public will get a chance to make its reaction known. Even in the current trend toward deregulation, no one has seriously suggested that transmission standards be left to the individual broadcasters or that spectrum assignments should no longer be made by the FCC. By setting standards and other ground rules, the Commission creates the environment in which the corporate entities that will provide service will function.

One good example of this kind of decision making is the support that the FCC gives to terrestrial broadcasters. Terrestrial broadcasting has immense support in Congress because it is the most used medium through which office holders get their message to voters. Many FCC regulations, such as the division of profits from reruns, appear to have been made with the primary purpose of keeping this industry alive.[13] Another example of regulation in the public interest, this time by act of Congress, was the All-Channel Receiver Act, which required all TV sets sold in the US to have UHF capability. This was a very successful example of government regulation of the free market that was to everyone's eventual benefit. Without it, many receivers would have been VHF-only, and the UHF spectrum would have proved impractical for TV.

2) The Need for High Spectrum Efficiency: NTSC has a very low spectrum efficiency. However, this is not due to stupidity on the part of its system designers. In 1941, when the standard originated, spectrum was not in short supply and cheap receivers had to have limited processing power. Neither of these conditions holds today. The electromagnetic spectrum is now a strictly limited natural resource. While the available spectrum is steadily being expanded at the upper end by advances in technology, TV occupies a large block of the more easily used UHF and VHF bands. In addition, it is now more practical to put a substantial amount of processing power into consumer products.

With the growth of mobile applications, pressure on the FCC to release unused UHF spectrum mounted. It was the fear of broadcasters that they might need more spectrum to compete with HDTV provided by alternative media that led to the current FCC inquiry that is working on HDTV standards. This has proved to be a very fruitful Inquiry, as it is leading to methods that are much more spectrum-efficient than NTSC. If the FCC's plan to turn off NTSC 15 years after HDTV broadcasting starts is actually carried out, we shall have at least the same amount of service as now within a considerably smaller spectrum allocation.

a) The role of source coding: It is obvious that if less bandwidth can be used for video of a given quality, or if quality can be improved without expanding bandwidth, the spectrum efficiency goes up. Until 1990 and the GI proposal, most executives in the TV industry thought that the first idea was impossible but the second might be accomplished. Of course, if one is true, the other must also be true, since these two statements are different ways of describing the same phenomenon, which is an increase in spectrum efficiency.

The method that has given the highest compression so far with manageable complexity, and is therefore used in all modern video coding systems, is the application of the discrete cosine transform (DCT) to the motion-compensated prediction error. Since provision must be made for scene changes and station switching, it is necessary to transmit some nondifferential information as well, either continuously (as the "leak" in DPCM) or from time to time. The net result is that the GA system can deal with no more than about three independent frames/s. For all its faults, uncoded NTSC can transmit 30 entirely independent frames each second, and each frame can comprise an arbitrary assemblage of sample values. The savings due to coding are dependent on successive frames being highly correlated and on each frame having high spatial autocorrelation (the efficiency of the DCT itself depends on the latter). While both of these situations are nearly always as stated, sometimes this will not be the case, and some new kinds of degradation will be evident [8].

b) The role of channel coding: One goal of channel coding is to fit as many programs as possible in each locality within the overall spectrum allocation for the service. This capability, although frequently ignored, is just as important as the compression achieved by source coding, which is universally recognized. In the US, at present, we can use about 20 channels in each locality out of 67 that are allocated, while in Britain the ratio is 4:44. Modern methods, as discussed below, may raise this ratio to 1:1. This would be just as important as reducing the bandwidth of a single program from 6 to 1.76 MHz!

The limitation of 20 out of 67; i.e., the existence of 47 "taboo" channels in each area, is due to a number of factors. The most fundamental, and hardest to deal with, is cochannel interference from another station on the same channel in an adjacent area. Given the carrier-to-interference ratio required for proper operation, the effective radiated power (ERP) of the transmitter, and the capability of a certain receiving antenna[14] and receiver, it is possible to calculate the minimum separation of stations, which is 160 mi for NTSC. This must be reduced to about 100 mi for HDTV in order to permit giving a second channel to each current broadcaster in accordance with the FCC's intended transition scenario. Clearly, HDTV must have much better interference performance than NTSC.

The second most important taboo is that adjacent channels cannot be used in the same cities, as discussed in Section II-A-2. The remaining taboos are predicated on poor receiver performance and are outdated. They need not apply to a new TV system.

[13] A topical example of government support for terrestrial broadcasting was the decision by the US Supreme Court on June 27, 1994, in which the economic viability of the broadcast industry was accepted as a legal basis for the reinstatement of the rule requiring cable companies to carry the local over-the-air programs. See L. Greenhouse, "Justices Back Cable Regulation," *NY Times*, June 28, 1994, p. D1.

[14] The antenna assumption is one of the "planning factors" established by the FCC to make it possible to calculate coverage area before a station goes on the air. The use of a better or worse antenna would make reception better or worse, but would not affect the calculation, which, to be useful, must be on a standardized basis.

3) Must All Receivers Have the Same Picture and Sound Quality? Television programs can be enjoyed over a wide range of image quality as long as the sound is free of serious distortion. At present, there is a wide variation of image quality from receiver to receiver. This is caused partly by differences in the size and quality of receivers and is also due to great variations of the amplitude and quality of received signals. The latter is affected by the kind of antenna used as well as by local conditions of signal strength, interference, and ghosts. These facts are widely recognized by the public as well as by TV professionals, although not often verbalized. No one, including the FCC, expects equally good pictures on all receivers; there is no FCC regulation of receiver image quality. On the contrary, should the FCC attempt to specify minimum receiver performance, there surely would be a storm of protest both from manufacturers and from free marketeers.

a) Receiver price versus performance: Typical households have two or three receivers. The best and largest is usually in the living room, while the others are in secondary locations such as the kitchen, children's rooms, etc. The latter, if bought for the purpose, are usually smaller and cheaper. While consumers certainly would not object to having maximum quality on all receivers, they have come to expect, as they do with most other products, that the cheaper sets will have lower performance. What would trouble consumers a good bit more would be the nonavailability of low-cost sets for these less critical uses.

In NTSC, it is possible for manufacturers to provide this range of price and performance because the main cost is the cabinet and display, compared to which the cost of the circuitry is almost negligible. This is not likely to be true with HDTV. Even in the largest and most expensive sets, signal processing will be an important part of the cost. If a complete decoder is required in all receivers, it will be the main cost in small sets. As long as this condition holds, it will not be possible to make inexpensive sets for today's less-critical applications.

This problem would be much less severe if simulcasting of NTSC were to remain in place indefinitely. However, the FCC's plan to take back a large proportion of the spectrum now allocated to TV requires abandonment of NTSC at some point. The lack of cheap receivers that can deal directly with the HDTV signal (or the lack of cheap set-top converters, which depend on the same technology) may prove an insurmountable obstacle to ever shutting NTSC down.

b) Portable and mobile receivers: While mobile receivers are not a big factor in the US, a very large proportion of sets in homes are portable in the sense that they may be moved from place to place and generally use on-set antennas—"rabbit ears." Well over half of the receivers in the US have antennas rather than being connected to cable or to satellite ground stations. This is a remarkable situation, since nearly two-thirds of TV homes in the US are on cable.[15] What makes these ratios important is that the coverage performance of proposed HDTV systems is predicated on the use of a properly installed receiving antenna with 10 dB gain and 14 dB front-to-back ratio. One knowledgeable critic has even stated that, beyond 35 mi from the transmitting antenna, reliable reception will require a low-noise amplifier mounted on the antenna mast [9].

Under these conditions, it is clear that the abandonment of NTSC simulcasting will create a very difficult problem. Reception with rabbit ears will become unreliable, and coverage will be drastically reduced for receivers that do not have the assumed high-performance antenna. This will make it very difficult to maintain coverage and to provide low-cost receivers thus creating another obstacle to the FCC transition scenario.

4) Interoperability: Although there had been little talk of interoperability—the easy interchange of video data between systems of different performance, different applications, different industries, and different vintages—before it was raised in a very forceful way by computer interests [10], the frequent need for transcoding makes interoperability of great importance within the TV industry itself. The FCC eventually recognized this need by making interoperability a subject to be discussed in the Inquiry.

a) The need within the TV industry: Considering the large number of standards now in use and the still-unsolved problem of converting between NTSC and PAL,[16] one would have thought that it would not need FCC oversight to guarantee that transcoding would be taken into account during the design of a new system. Yet this was not the case. For example, the NHK system, which was the first format proposed for use as an international exchange standard, has scan rates that make it difficult to transcode either to PAL or NTSC.

The discussion in Section II-B-3 about the need for receivers with different price and performance illustrates that interoperability is not just a burden placed on the TV industry for the benefit of the computer industry, as is often stated. The ability to make simple receivers that can deal with a complex signal, even if their image quality is not as good as that of expensive receivers, is the key ability that is needed. It is so fundamental to system design that it cannot be added at a later date.

b) Nondisruptive improvement over time: Even before the computer industry was calling for an HDTV system that could easily be handled by workstations, the FCC itself was calling for "nondisruptive improvement over time." Learning from the NTSC experience, the Commission has

[15] These numbers are estimated from data provided by the Cable Advertising Bureau and Paul Kagan Assoc. Data from NCTA and Nielsen was also consulted.

[16] In spite of long effort, today's best transcoders are far from perfect, as was clearly demonstrated during the 1992 Summer Olympics. This event was shot in PAL and converted to NTSC for airing in the US. Defective rendition of rapid motion, such as disappearing volleyballs, was obvious, even though it detracted little from the popularity of the broadcasts. The reason transcoding is so hard is probably the prevalence of a great deal of temporal aliasing in all current TV systems.

made plain that any new system ought to be able to be upgraded without making earlier receivers obsolete. NTSC has very little room for progress in this way. The main change made since color was added in 1953 was stereo audio.[17] Any improvement in picture quality since 1941 is due to better cameras and picture tubes, and not to any change in system standards.

It does not take much reflection to show that, to improve the quality of a system after installation, it is necessary to send additional data that only new receivers would use. This data either must be hidden within the existing signal in such as way as not to degrade image quality on existing receivers or must be transmitted in a separate channel. In either case, many defects of the original system will remain in the enhanced system, even in new receivers.[18]

c) Across applications and industries: Interoperability became a public issue when it became apparent to the computer industry that the ability to display good-quality video on computer screens was very important to the future of the industry. With the still-declining cost of processing power, revenues can be kept up only by increasing the amount of computation. Nothing is so computation-intensive as high-resolution moving images. Even today's computers have a video screen, and many of the multimedia applications coming into use depend very heavily on video. It seems quite natural, therefore to display broadcast video on computers and to use computers to generate video sequences.[19]

Another industry that is affected is electronic imaging. Although no one thinks that film is going to disappear in the near future, it has become quite feasible to handle high-quality imagery in electronic form for virtually any application. Amateur photography is a good example. While equipment of full photographic quality is still too expensive for most users, properly handled images having a real resolution of 500–1000 lines are acceptable in many cases. If HDTV frames could be used as snapshots, an entire industry might be created. Similar possibilities exist in medical care, education, and publishing. The minimum demand of these non-TV industries is progressive scan and "square pixels." (equal horizontal and vertical resolution) What the TV industry is so far willing to give is all-digital transmission plus a self-description of each transmission by means of embedded headers and descriptors.[20]

[17]Since the addition of color substantially reduced the luminance resolution of receivers, existing or to be manufactured, and added cross color and cross luminance to the jargon, one would have to say that the 1953 changes, while praiseworthy, were not entirely "compatible."

[18]The extreme vulnerability of NTSC to interference and the associated poor spectrum efficiency as well as all the disadvantages of interlace, are related to its system design and cannot be cured by improved receivers. Ghost cancellers might well improve the performance of new NTSC receivers. The system described in Section III-D is specifically designed to permit upgrading over time.

[19]Computers are already widely used to create and edit video in the NTSC format. Unwieldy as it is, it has nevertheless proved quite feasible to design the hardware and software needed for this application.

[20]The TV industry is not a monolith on this or any other question. For example, ABC and FOX, two of the four TV networks, favor progressive scan.

C. The Transmission of Media and Their Characteristics

In the US, at present, video signals are transmitted to receivers by means of terrestrial (over-the-air) transmission, by cable, by VCR, and by satellite. Each of these media has different physical characteristics that must be taken into account in order to get the best results. The last is by far the least important in the US, since it is confined to a few million users who tune in directly on the programs being sent to TV stations and to cable head ends. However, this year a satellite has been launched and two operators, DirecTV and USSB, are providing service. Initial acceptance has been good, so the situation may change.

1) Terrestrial Transmission: Terrestrial transmission is the most popular medium in terms of receivers served. It is free in the US and widely used for political purposes, giving it immense support from the public and in Congress. Technically, it is the worst medium, suffering from noise, ghosts, interference, and frequency distortion. A unique characteristic is the very wide variation in signal strength from receiver to receiver. Coupled with the differences in receiver noise performance and antenna characteristics, a very wide variation in CNR is encountered, corresponding to more than a 5:1 range of channel capacity. The NTSC signal design is such, however, that good synchronization and good audio quality are maintained under virtually all conditions in which the image is even marginally viewable. Very simple antennas can be used except at the boundary of the service area. In the absence of interference, with a good receiving antenna, and with a line of sight to the transmitting antennas, programs can be viewed some 200 mi from the transmitter site.

Twelve VHF and 55 UHF channels are allocated for TV, with a maximum of seven VHF and about 12 UHF stations actually licensed in each city.[21] Adjacent channels are not used in any one locality and stations on the same channel must be at least 160 mi apart. Broadcasters greatly prefer VHF assignments, since better coverage is obtained with lower transmitter power. In the absence of cochannel interference, and using the maximum permitted ERP, coverage is noise-limited somewhat beyond the radio horizon—52 mi for an antenna 1350 ft above the ground (HAAT). In certain areas of the country, HAAT's of as much as 2000 feet may be used. This has a radio horizon of 63 mi, but a noise-limited range of 80 (channel 2) to 67 (channel 69) mi. Actually, few stations have maximum-height antennas.[22]

2) Cable: Cable service is available to about 96% of the 95 million TV homes in the US and about 65% actually subscribe. Although cable provides a much larger number of programs than terrestrial broadcasting, most cable viewing is of programs that originate with the networks. In principle, all of the technical problems mentioned in connection with over-the-air transmission ought to be absent on cable, but they are not.

[21]On average, each television household in the US has 13.3 free stations available to it (Nielsen).

[22]Information on antenna heights from Dr. T. J. Vaughn of Micro-Communications, Inc.

At present, cable uses trunk-and-branch distribution, with amplifiers along the trunks as needed. Some nonlinear distortion is introduced in this way. Coaxial cable is almost always used into the residence, but fiber is steadily replacing cable on the trunks. Cable is not completely impervious to leakage either in or out, so the same kind of natural and man-made noise is encountered as in terrestrial broadcasting, although to a lesser degree. Passive lossy signal splitters are used in many locations, with unused taps generally unterminated. This creates a kind of endemic multipath that behaves much like a low-pass filter.[23]

Signal strength from receiver to receiver is more uniform than over the air, but still varies because of the use of signal splitters. All channels have signals of about equal amplitude, so that there is no adjacent-channel taboo as in terrestrial. Cable companies try to ensure 38–40 dB CNR at the receiver terminals, but do not always succeed. If they did, the noise would be marginally visible but not annoying. In spite of all this, "cable quality" is generally superior to average quality with rabbit ears. In many locations, however, a good antenna produces better quality than provided by cable. Informed opinion is that viewers usually subscribe to cable because of a wider choice of programs, and not for higher image quality.

3) Video Recorders: For every two receivers in American homes, there is one VCR.[24] Although originally used mainly for time-shifting, the vast majority are now used for playing rented movies.[25] There are also about 22 million camcorders. Thus, tape viewing accounts for a significant portion of TV use. Any new system *must* have affordable and reliable VCR's to be acceptable.

Getting two hours of NTSC signal onto a small spool of tape was a remarkable technological achievement that required some compromises with signal quality. Sometimes, "VHS quality" is used as a measure to indicate something considerably below that of NTSC. Certainly, the resolution and SNR of the VHS format is lower that of studio-quality NTSC. However, NTSC as typically viewed in the home is also quite inferior to NTSC in the studio. My own opinion is that with a good tape and a VCR in good condition, one gets better pictures, on average, from tape than from broadcasts.

4) Satellite Broadcasting: In principle and in practice, the satellite channel is substantially superior to all other existing means of transmitting video to the home. A line-of-sight path is always used, along with directional antennas. There is very little multipath and little adjacent-channel interference. Cochannel interference would be much like that of terrestrial broadcasting from a single centralized antenna. Most current transmission, which was never intended for broadcasting, is analog FM using an RF bandwidth of 36 or 54 MHz. This gives a favorable "triangular" noise spectrum. Some digital transmission is also used with a very conservative data rate of only 45 Mb/s. The system noise budget is arranged so that even under extreme weather condition such as heavy rainstorms, the received signal is well above the threshold, and reception is studio quality.

For DBS to the home, a bandwidth of 24 MHz will be used. For the less demanding requirements of home reception, it will most likely be found that a gross data rate of some 60 Mb/s per channel can be used as compared with 20–25 Mb/s for terrestrial broadcasting. This will permit transmission of two HDTV signals or 8 standard-resolution signals, with far higher reliability than is likely to be experienced with terrestrial transmission.

III. Some Possible Solutions

A. Source and Channel Coding

Shannon's work can be interpreted to mean that source and channel coding ought to be independent. In this approach, the source coder removes all statistical redundancy, producing a signal that looks like random noise; the channel coder adds redundancy in just the right way so as to permit near-perfect error correction. Each coded bit is then essential to reconstruction. However, such a scheme is impossible to implement exactly, since all redundancy cannot be removed. If it were, a single error would make further decoding and resynchronization impossible. The closer we get to such an "ideal" system, the more fragile the signal, the longer the coding and decoding delays, and the more difficult the synchronization.

In the best current systems, the data transmitted is very far from being equally important. In addition, the concept applies only to point-to-point systems in which the receiver CNR is well defined. It does not apply to broadcasting, in which very large differences in CNR are found from receiver to receiver.[26] Thus terrestrial broadcasting requires a rethinking of the coding problem if optimum use is to be made of the limited spectrum that is available.

There are two approaches that can be taken. Using high-power centralized transmitters as at present, one solution involves self-optimization at each receiver according to the amount of data that can be recovered. The latter should be as close as possible to the Shannon capacity at that receiver. Necessarily, everyone does not get images of equal quality. The second solution involves making the signal strength, and therefore the channel capacity, as nearly uniform as possible across the population of receivers. This can be done by using a cellular network of low-power transmitters, all emitting the same program. If the transmitters in the cellular network all operate on the same frequency, the arrangement is called a single-frequency network (SFN). The receiving area can be delineated almost arbitrarily by the placement of the transmitters, and contiguous areas can use the same channel for different programs. This

[23] In the US, it is not unusual to find ghosts on cable similar to those encountered in over-the-air reception. In most cases, these ghosts were present in the signal when received at the cable head end.

[24] Data from Zenith Electronics Corporation.

[25] This may not be true in other countries, where time-shifting is a common practice.

[26] The broadcasting problem, unfortunately, has attracted very little attention from the theorists [11].

method achieves the highest possible spectrum efficiency; cochannel interference disappears as a design issue. Only as many channels need be allocated to TV service as the number of independent programs that are to be available in each locality.

1) Multiresolution by Combined Source and Channel Coding: In analog systems, image quality necessarily deteriorates steadily with falling signal quality, primarily through lower SNR. The resulting soft threshold can be thought of as a rough kind of self-optimization (The sound quality remains good at a signal level that produces barely watchable images, and that is probably a good choice to make in new systems). To achieve the very high compression ratio needed to transmit HDTV in a 6-MHz channel, at least some digital data must be transmitted. In digital transmission, there are no known methods of getting a soft threshold, i.e., of recovering a continuously higher digital data rate from a continuously rising CNR.[27] Thus recovery must be a stepwise affair. This means that the source coder must organize its output into a number of data streams in which the quality increases with the number of streams recovered. The channel coder must package these data streams in the transmitted signal in such a way that the number of streams recovered increases in a stepwise fashion with receiver performance and with the signal strength at the receiver terminals. Finally, the receiver must make the best possible picture from the recovered data at each level of CNR.

Resolution and SNR are the two image-quality factors that depend on the amount of data recovered. There is no consensus as to which should be varied the most from level to level; MPEG2 provides both possibilities [12]. A small amount of white or high-frequency noise is relatively harmless, but an amount and character of noise much different from what is now seen when reception is deemed acceptable is probably unwise. On the other hand, there is clearly a very large tolerance for resolution differences, as today's situation makes obvious. This is not only true for small receivers, which look sharp even when the resolution in absolute terms (number of samples per picture dimension) is quite low. It is also true for large displays. Their resolution in absolute terms is quite low, but they are nevertheless preferred. In audience tests at MIT, image size was by far the most important factor in viewer preference [13]. Viewing angle, which is of great importance in subjective assessment of TV displays, cannot be controlled by the system designer.

These observations provide enough direction for designing a system using several levels of quality. We shall designate such systems as using multiresolution (MR) coding as distinct from single-resolution (SR) coding, even though both resolution and SNR may vary from level to level.

[27] In [38], the authors describe a spread-spectrum method that produces a quasi-continuous threshold for the channel coder. It is not clear whether adding transform coefficients in a quasi-continuous manner will give good picture quality at all levels.

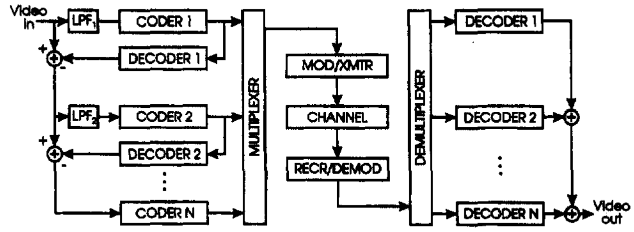


Fig. 1. *Pyramid Coding.* This is the basic arrangement of a multiresolution system that provides good picture quality at every level of performance. A low-pass filter (2- or 3-d) selects information that is to be included at the lowest-quality level. This is coded and decoded and then subtracted from the original video. A second low-pass filter provides information for the next (enhancement) level, which is also coded, decoded and subtracted from the remaining input video, etc. (Subtracting decoded data at each level ensures that any coding distortion is available to the next higher level for possible correction.) The coded data streams from all the levels are multiplexed, modulated, and transmitted. The receiver combines the decoded lowest level with whichever enhancement levels are recovered to produce the best picture that can be made from the available data.

a) Multiresolution source coding: There is a considerable literature on MR systems, as they are useful in a number of applications, including browsing through image data bases.[28] An early paper coined the term "pyramid coding" for schemes in which a basic image could be upgraded by addition of more information, as shown in Fig. 1 [15]. The general idea was used in a number of proposed receiver-compatible HDTV systems for the US in which enhancement data, either hidden within the main signal or transmitted in a second channel, would be added to a standard NTSC signal [16].

A significant aspect of pyramid coding is that, to be useful, all the pictures in the hierarchy must be free of obvious defects such as ringing (Gibbs phenomenon) due to sharp-cutting filters. To avoid this problem, the filters that separate the several data streams must have a smooth and not-too-rapid cutoff. As a result, the same frequency component may be represented in more than one stream. With existing coding technology, this results in a penalty in the quality/compression tradeoff as compared with systems that code the entire image spectrum in one stream. In general, pyramid systems require a somewhat higher data rate at their highest level to achieve the same quality as that of SR systems. This is offset by the ability of MR systems to provide good pictures, albeit of lower resolution, at lower data rates which permit greater coverage. MR systems can also provide higher quality than SR systems when it is possible to deliver more data to the decoder.

b) Multiresolution channel coding: For digital transmission, it is sometimes suggested that unequal error protection can be used to achieve multiresolution [17]. However, the numbers do not work out very well. The amount of error protection required at low CNR is very large and leaves little room for the real data. Another proposal is to subdivide the channel by frequency or time, using constella-

[28] This was sometimes called "progressive transmission," which must be carefully distinguished from progressive scanning [14].

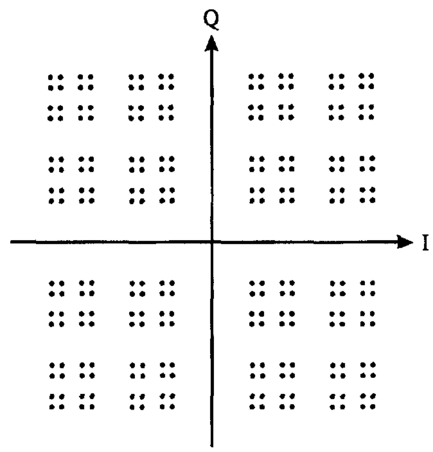


Fig. 2. *Nonuniform Constellation.* This constellation has four levels of performance with CNR thresholds approximately 6 dB apart. It is intended to be used with a multiresolution source-coding method that produces four streams of data.

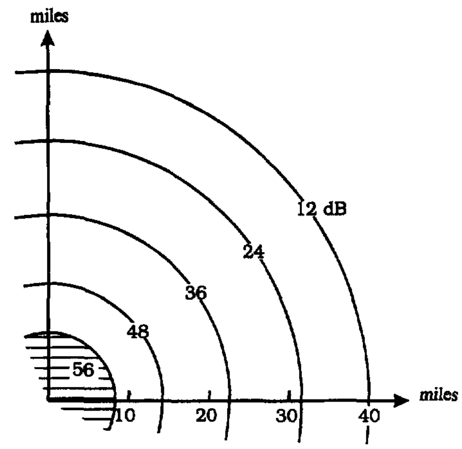


Fig. 3. *Variation of CNR With Range.* The inverse-square law does not govern typical TV antenna performance. This is because of its height and the vertical profile of its beam, as well as high attenuation at the edge of the service area. Grazing incidence in this area causes the field strength to diminish very rapidly with distance. Finally, the FCC planning factors, which rise with distance, effectively reduce the field strength, producing the result shown. The most notable features are the near-uniform field strength in the inner 8 mi and the uniform decrease in signal (in dB) with distance. Note that the channel capacity, which is proportional to CNR in dB, is more than four times as high downtown as at the threshold of service. (Data from Dr. O. Bendov.)

tions of different density (different numbers of bits/cycle) in the various subchannels. This is also inefficient, since at the threshold CNR for a dense constellation (i.e., finely quantized), subchannels with less dense constellations (coarsely quantized) are very inefficient. At the present time, the best known method is to use a multilevel modulation scheme such as the nonuniform constellation as in Fig. 2.

As is the case with MR source coding, MR channel coding is also somewhat less efficient than SR coding at the design threshold of the latter. However, the MR system becomes more efficient than the SR system at higher CNR. In addition, the former can deliver pictures, albeit of lower quality than that of the latter, at substantially lower CNR, thus extending the coverage area.

c) Overall performance of MR systems: The variation of receiver CNR with range for a typical current-day UHF transmitting antenna is shown in Fig. 3.[29] Note that the channel capacity, which is proportional to the CNR, decreases by a factor of more than four from the central to the outlying area. Obviously, sending the same data rate to all receivers wastes a great deal of capacity in just those close-in areas where spectrum is in shortest supply.

In Fig. 4, a comparison is made between the performance of SR and MR systems, in which the design threshold of the former is 16 dB. In such an SR system, an HDTV image of uniform quality is delivered everywhere the CNR is at least 16 dB, and no picture at all is delivered beyond that. In the MR system shown, a low-resolution image is delivered from 6 to 16 dB, a medium resolution image at 16–26 dB, an HDTV image similar to that of the SR system at 26–36 dB, and a better-than-HDTV image for CNR's in excess of 36 dB. In qualitative terms, the MR scheme extends the service area considerably beyond that of the SR system and delivers superior pictures for CNR's higher than 36 dB. The price paid is a reduction in quality for CNR's between 16 and 26 dB. While these numbers are not associated with any particular system, they are believed to be typical.

2) Single-Frequency Networks: Although the SFN concept is not new, it was recently brought to prominence by its proposed use in digital audio broadcasting in Europe [18]. It is also used in some radio applications [19]. The entire service area of a station can be covered with a cellular array of same-frequency low-power transmitters, or the array can be used in the outer region and a single medium-power transmitter, or even a satellite broadcast, can be used for the central region. The various transmitters may be fed by cable or in a different channel, or all transmitters may derive their signals from each other. The carriers may be identical or intentionally offset. Some successful field tests have been carried out, but no full-scale SFN has yet been implemented. There is considerable controversy over details of the expected performance [20].

Within the cellular array, the signals from a group of nearby transmitters appear as multipath at the receiver. The amount of multipath can be reduced, but not eliminated, by use of directional antennas [21], but it would be far preferable to use simple antennas, perhaps omnidirectional, in a large percentage of locations. Thus the multipath performance of the modulation and channel-coding system emerges as a principal concern. Multipath is a linear distortion, equivalent to the effect of a certain filter. Its two main effects are intersymbol interference (ISI) and a possible increase in noise level due to equalization of the multipath distortion.

[29] This diagram takes account of the "planning factors" used by the FCC in determining coverage. Among other things, these factors deal with the percentage of times and percentages of locations in which the given reception conditions are met or exceeded. In the central area, signal strength is nearly constant. This is due to the vertical profile of the transmitter's antenna beam and to the fact that the receiving antennas are much closer to the ground.

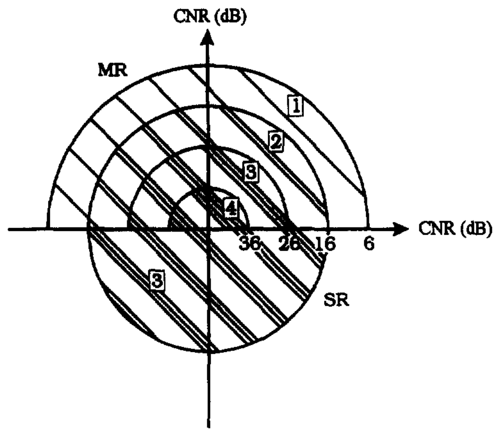


Fig. 4. *Comparison of Typical Single- and Multiple-Resolution Systems.* The thresholds are shown in circle and the quality levels in squares. The SR system provides the 3rd level of quality everywhere where the CNR is >16 dB. The MR system provides a larger service area (out to 6 dB) at lower quality (1st level) and higher-than-SR quality (4th level) where the CNR >36 dB. The price for this improved overall performance is lower quality (2nd level) between 16 and 26 dB. The two systems have the same quality from 26 to 36 dB. The numbers here do not represent any particular MR system; they are intended to show a typical relationship between the service rendered by an MR and a SR system using compression schemes of roughly equal effectiveness.

While the main advantage of SFN's is spectrum efficiency, there are other advantages as well. Service areas can be of irregular shape, and can include regions that are otherwise denied reception because of intervening obstructions. Except for a narrow region along the boundary of the service area, the transmission power can be raised enough so that CNR is no longer a factor in reception. Even so, the total emitted power is much less than that needed by a single centralized transmitter.[30] Note that the improvement in spectrum efficiency due to MR coding is less important in SFN's than in the conventional single-transmitter arrangement. However, the facilitation of the manufacture of receivers of a range of price and performance makes MR coding advantageous in all cases.

ISI due to multipath reception can be removed by equalization or by use of multicarrier reception as discussed below. The accuracy, complexity, and noise performance of these schemes are the main issues.

3) Multicarrier Modulation: The distorting effect (the ISI) produced by a given level of multipath depends not only on the total power and relative delay of the echoes but also on the ratio of the temporal spread of the echoes to the symbol length of the signal. In VHF and UHF terrestrial transmission, most echoes occur within about 20 μs of the main signal. This does not cause much trouble with AM or FM audio broadcasting, with a symbol length of about 25 μs, but it produces heavy impairment in television, with a symbol length of about 120 ns. Obviously, one way to reduce (but not eliminate) the distortion is to divide the signal into a large number of components, each of which has a much longer symbol length, and to transmit these components as narrowband modulated carriers within the original channel. The ISI can be eliminated completely by inserting after each symbol a guard interval during which a portion of the symbol waveform is replicated. This permits integrating each symbol over its symbol duration without unintentionally including energy from symbols just before or after the symbol being demodulated. The guard interval itself must be longer than the multipath spread. Since the guard interval reduces the efficiency of the transmission, it is advisable to make the symbol long as compared with the guard interval, with a correspondingly large number of carriers.

Frequency-division multiplex, as discussed above, has been improved by two developments—orthogonalization of the modulated carriers so that no bandwidth need be wasted by using guard bands, and implementation by means of the discrete Fourier transform [22]. The resulting system, including coding, is called coded orthogonal frequency-division multiplex (COFDM). It is already used in some modems for digital data transmission over telephone lines, and is being planned for use in digital audio broadcasting in Europe [23]. It is the subject of a companion paper in this journal [24].

Another important property of OFDM is that out-of-band radiation is much less than in single-carrier modulation (SCM). This is because orthogonality, as produced by the discrete Fourier transform (DFT), makes the spectrum of each modulated carrier have the shape $(\sin(\omega)/\omega)$ centered on the carrier frequency, with the zeroes placed at the locations of the neighboring carriers. With hundreds, or even thousands, of carriers, the spectrum thus decays extremely rapidly at the edge of the channel, even without filters.

The elimination of ISI by OFDM, although very valuable, is not a complete solution to the transmission problem, as we must still deal with the noise caused by equalization of the multipath channel. Originally, the claim was made that COFDM adds echo power constructively, so that the error rate actually goes down with more echoes. While it is true that, averaged over all receivers, the powers of signal and ghosts do add, this is not true at every individual receiver (The BER goes down in some cases and up in others). Depending on the precise character of the echoes, deep notches may be produced in the spectrum. The worst case is that of a single echo of 0 dB, which produces actual nulls. Data transmitted on carriers at frequencies where the signal strength is very small is obviously less reliable. This can be dealt with by interleaving and coding, but it is clear that, at some locations, transmission may be adversely affected. One remedy is the use of directional antennas at those locations. In most cases, these would not have to be very elaborate, as it is only necessary to reduce the offending ghost by 3–6 dB. Simple dipoles would suffice in many cases.

Wideband nulls can also be caused by radio-frequency phase cancellation. A solution in most such cases of this

[30] Single transmitters are remarkably inefficient in covering large areas on account of the very rapid decrease in signal strength with distance near the boundary of the service area. It takes an increase in power of from 1 to 1.5 dB to increase the range by 1 mi.

kind is simply to move the antenna by a fraction of a wavelength. More elaborate installations could use space-diversity reception.[31]

The tradeoff in complexity between receivers for SCM and for OFDM involves the time-domain equalizer used in the former versus the DFT required for the latter. In OFDM, a frequency-domain equalizer, which is far simpler than a time-domain equalizer, is most natural. On the other hand, OFDM requires the DFT operation, which is not needed in SCM.

4) Digital versus Hybrid Transmission: In the "ideal" system discussed in Section III-D, we use hybrid analog/digital transmission. This undoubtedly seems a quaint idea from the past to those who have joined the digital bandwagon. However, careful analysis of some specific aspects of coding systems shows that digital transmission does not have all the advantages claimed for it. It is true that some digital data must be transmitted in order to achieve the very high compression associated with motion-compensated transform coding. However, it is also true that higher channel-coding efficiency can be achieved with hybrid transmission. Finally, interoperability is not materially enhanced by all-digital transmission.

a) Source-coding efficiency: In motion-compensated transform coding, the amplitudes and identification of adaptively selected transform coefficients comprise the bulk of the data to be transmitted. In the GA system, this data is jointly coded for 2–3 million coefficients per second at about 4–6 b/coefficient. In fact, the nature of the large correlation between amplitude and identification (the spatial frequency of each selected coefficient) is such that not much would be lost by separately coding the two kinds of data. (This is discussed further in Section III-D-1.) If the statistical relationship among the coefficient amplitudes themselves is not utilized in the coding scheme, there is nothing to be gained by quantizing the amplitudes before transmission. That simply adds quantization noise. Analog transmission works well in this case. The data that must be transmitted per coefficient in a hybrid system is one analog sample plus less than one bit. All other aspects of MPEG coding can be used with hybrid analog/digital transmission, so that comparable compression ratios can be achieved.

b) Channel-coding efficiency: In Section II-A-2, we pointed out that, when analog information (such as the amplitude of transform coefficients) is sampled and quantized for digital transmission in an analog channel, the requirements for achieving a transmission rate close to the Shannon rate include very fine quantization combined with very effective error correction. Note that noise added to these coefficients produces no catastrophe in the reconstructed image; thus, they need not be entirely noise- and error-free. The requirement for near-perfect transmission in MPEG-like systems arises from joint coding of the amplitudes with the adaptive-selection data, for which errors produce serious image defects. On the other hand, analog transmission of the coefficient amplitudes can readily achieve the full Shannon capacity, and it can do this for a range of CNR, and not only the threshold CNR. For the peak-power-limited additive-white-noise analog channel, if the coefficients comprise a train of uncorrelated analog samples of uniform amplitude probability distribution, the mutual information (i.e., what the noisy output signal tells us about the noiseless input signal) is equal to the Shannon capacity of the analog channel in which they are transmitted. (For an RMS-power-limited channel, a Gaussian distribution is optimum.)

Since the coefficients to be coded represent differential data, i.e., prediction error, and must therefore be integrated to generate the desired output, it may be thought that analog transmission cannot be used because of the possibility of a catastrophic accumulation of noise in the decoder output. The coefficients in their analog form have precisely zero average value, as does the channel noise. The average is approached fast enough so that no catastrophe occurs, as we have demonstrated in our simulation. The "integrator" in this case can have zero response at zero frequency and still produce the desired output.

c) Interoperability: The difficulty of transcoding between two different video signals is primarily a function of their relative sampling grids. It makes little difference if the signals are in digital or analog form, since conversion from one form to the other is rather simple. If the signals are compressed, it is generally necessary to convert to uncoded form to do any transcoding at all.

The fact that the two systems have different spatial sampling frequencies does not present much of a problem since the sampling theorem provides the theoretical basis for moving from one grid to another. In practice, filters should be chosen with due regard for perceptual effects [25]. Different temporal sampling rates, however, always cause trouble. This is because temporal aliasing is nearly always present unless motion is less than one sample/frame. The aliasing greatly inhibits temporal filtering, which is prone to produce defects such as multiple images. With the amount of motion commonly encountered, a rate of even hundreds of frames/s is insufficient to allow the elimination of temporal aliasing without excessive blurring. Blurring of moving objects is counted as a defect to such an extent that electronic shutters are sometimes used although this makes the aliasing worse.

Good temporal interpolation can only be done if motion compensation is used. While this is quite complex, good results can be achieved. In Ph.D. dissertations by Martinez and Krause [26], essentially flawless transcoding was demonstrated with arbitrary ratios of frame rates.

Another factor in interoperability is the complexity of the relationship between the transmitted signal and the uncoded video signal that it represents. *High compression*

[31] A single echo causes the frequency response to undulate over the band with a frequency separation between peaks equal to the reciprocal of the relative delays. If the relative delay is comparable to the reciprocal of the radio-frequency (RF) bandwidth, a single cycle of the undulation is about as wide as the rf band. Assuming that the signals come from different directions, the null can then be moved a great deal by shifting the antenna on the order of one wavelength. In general, the antenna has to be moved on the order of the velocity of light (c) multiplied by the relative delay. The exact amount depends on the directions of the signals. For relative delays of more than λ/c, antenna diversity is not practical.

ratios necessarily involve complex coding algorithms. If it is necessary to decode an HDTV transmission completely in order to extract a low-resolution video signal for display in a small low-performance receiver, the receiver cannot be so low in cost. It is much better to use a pyramid coding scheme in which the simplest receivers deal only with the lowest level of the pyramid and can therefore use the simplest and least expensive decoder.

Interoperability is also affected by the channel coding scheme. Ideally, one would like a range of encoders of different quality (resolution) to be able to communicate with a range of decoders. In this way receivers of different price and performance could all accept the same transmitted signal, while the signals transmitted from a range of encoders of different resolution would all be acceptable by all decoders. One way in which this can be done is discussed in Section III-D.

B. Noise and Interference Control

Noise can usually be defeated by transmitting at higher power, although some limits are set by practical and economic considerations. However, the main limitation on transmitted power comes from the need not to interfere excessively with other stations. In the case of HDTV, the FCC's intended transition scenario calls for adding HDTV stations while current NTSC stations remain on the air. This must be done without materially reducing the latter's coverage, while at the same time attaining adequate coverage for the new transmissions. After NTSC is shut down, only HDTV stations will remain on the air, and they must have coverage similar to today's stations, but within a reduced overall spectrum allocation. It is clear that HDTV signals must be recoverable at lower CNR than now required for NTSC and that they must have better interference performance. To the extent that digital data is transmitted, error correction and concealment must be implemented in order to achieve appropriate image and sound quality. To the extent that analog information is transmitted, the recovered signals must have appropriate SNR.

For best noise performance in the additive white Gaussian noise channel, the spectrum of signals should be uniform.

1) Noise Performance for Digital Data: Within a given channel capacity as limited by bandwidth and CNR, errors caused by noise are correctable, in principle, by coding, as long as the Shannon rate is not exceeded. The closer the total transmission rate (signal data plus error-correction data) to the Shannon channel capacity, the higher the uncorrected (raw) error rate. To achieve net transmission rates that are a substantial fraction of the Shannon rate, the raw error rate must be quite high. A combination of outer Reed/Solomon plus inner trellis coding has proved to be an effective method with manageable complexity and coding delay [27]. A corrected bit-error rate (BER) of 5×10^6 is the generally accepted threshold of service, as error concealment is effective at that rate.

All digital modulation methods have sharper thresholds than analog schemes, and coded digital methods have extremely sharp thresholds. In analog systems, which have soft thresholds, coverage is usually calculated on the basis of a CNR that is exceeded in half the homes half of the time. There is as yet no generally agreed-upon values for these percentages for digital transmission, but it is clear that reception must be guaranteed much more than 50% of the time.

2) Noise Performance for Analog Data: In uncoded analog systems such as NTSC, the SNR of the recovered video signal is exactly equal to the CNR of the transmitted signal. In coded analog systems, such as FM or spread spectrum, it is possible to trade off bandwidth and SNR, although the tradeoff is generally not as effective as in digital modulation such as PCM. If the bandwidth of the data to be transmitted is less than that of the channel, an improvement in SNR can be achieved. For example, if 5 MHz is the usable channel bandwidth, 10^7 samples can be transmitted per second. If the number of samples to be transmitted is less than this, the SNR of the recovered signal can be higher than the channel CNR. With spread spectrum, if the different original signal samples require different SNR, then another improvement is possible by transmitting the more sensitive samples at relatively higher power without changing the statistical parameters of the signal in the channel [39].

3) Interference Performance: For a given relative power, analog signals interfere the least with each other when they appear to be random noise to each other.[32] This is easily accomplished with digital transmission, and is one of its major advantages, but rarely mentioned. One result is that the threshold carrier-to-noise ratio is about the same as the threshold carrier-to-interference ratio (CIR). Analog signals must be scrambled to accomplish the same end, and this is also readily accomplished with modern technology.

During the transition period to all-HDTV broadcasting, the interference between HDTV and NTSC is an important consideration. Interference is mutual; If A is less interfered with by B, it can be transmitted at lower power, thus interfering less with B. Of course, reducing power may reduce coverage where it is noise limited. It is much easier to plan the location and power levels of transmitters when no stations are already on the air in the band in question. When adding HDTV stations in the spectrum now allocated to NTSC, the problem is much more difficult. However, strong resistance to noise and interference is always helpful.

4) Synchronization and Accurate Carrier Recovery: Although not a factor in spectrum efficiency, synchronization of all clocks is a very important practical consideration. Accurate clock recovery is vital to minimizing the BER. The ability to synchronize rapidly and accurately in the presence of noise, multipath, and interference is essential to achieving proper coverage and is a great convenience when changing channels. One of the merits of NTSC is its ability to synchronize under very noisy conditions, a merit

[32] This is one of the most serious limitations of NTSC. Relative randomization of the scanning patterns would have greatly improved the interference performance. On the other hand, the known nonuniform spectrum of NTSC can be used to decrease its interference into fully randomized signals [28].

that is not surprising since more than 15% of the channel capacity is devoted to this purpose.

In principle, synchronization does not require the use of any channel capacity. If the system is well designed, statistical parameters of the signal, such as RMS value, autocorrelation function, etc., are well determined and can be used for this purpose. The use of synchronization signals not only uses some channel capacity, but inserts some periodicity into the signal, which increases its potential for interference with other signals. As a practical matter, and in view of the current state of the art, it appears than devoting a small amount of channel capacity to this function and accepting a slight increase in interference are defensible decisions. In the GA competition for the channel-coding scheme, the Zenith system, which does use pilot carriers, was able to synchronize at substantially lower CNR than the GI scheme, which did not. This was an important factor in choosing the former over the latter [31].

C. Multipath and Frequency Distortion Control

Multipath, which is a linear distortion, can be corrected by linear equalizing filters in the same manner as other sources of frequency distortion. Noise limits the performance of equalizers in two ways. If the uncorrected signal is noisy, calculation of the filter parameters must be done slowly enough so as to average out the noise. Even if the filter parameters are correct in terms of frequency response, a large increase in noise may result if there are near-nulls in the uncorrected spectrum. For SCM, errors are caused both by incompletely corrected frequency response, which leads to an imperfect "eye" pattern, or by noise, which also partially closes the eyes.

Echoes can be reduced in amplitude, but generally not completely removed, by use of highly directional receiving antennas. Almost whatever modulation and error-correction systems are used, it probably will always be necessary to use directional antennas at those locations that otherwise would have near nulls in the spectrum.

The situation is somewhat different in multicarrier modulation (MCM) because the data on carriers received at relatively low amplitude has a higher BER than data on carriers received at relatively high amplitude. The data in each transmitted block can be distributed across many carriers (preferably all of them) and the performance linked by a code. For example, the portion of the data with lower CNR can be weighted less heavily by the decoder [30].

There is very little data available on the effect of equalization on CNR in typical broadcasting situations. Recent tests at the Advanced Television Test Center using seven different combinations of echoes with a total power 7.5 dB below the direct signal have shown that the threshold CNR goes up, averaged over the seven echo sets, about 2.5 dB [31]. It should be kept in mind that much worse echoes are often encountered and that, therefore, a substantial reduction in coverage is likely if there are large echoes near the boundary of the service area.

1) Implementation of the Equalizer: Equalization can be carried out in the time domain or the frequency domain. In the time domain, an FIR filter somewhat longer than the temporal spread of the echoes is effective in most cases. The output is a linear combination of the signals at the various taps of the filter—typically 256 to 1024. The tap coefficients are obtained by various methods. Sometimes clock recovery is combined with coefficient calculation. Some methods use transmitted reference signals and some ("blind deconvolution") use the main received signal itself as reference [32].

In the frequency domain, equalization can be accomplished by dividing the channel output into a large number of narrow-band components and multiplying each by a single complex factor. This method is based on the assumption that the frequency response is constant across each narrow band, which is almost certainly justified when there are many hundreds of channels. The effect of such an equalization is exactly the same as that of a corresponding linear filter operating in the time domain. Note that in this form of equalization, a convenient pilot signal consists of an assemblage of sine waves or a swept-frequency signal, sometimes called a chirp. A convenient pilot signal for time-domain operation is one that determines the impulse response of the channel, such as a pulse.

Obviously, time-domain equalization is more natural for SCM and frequency-domain correction, which generally is much easier to implement, is more natural for MCM. However, there is no theoretical objection to interchanging these techniques, since the signal can be shifted easily, although at some expense, from one domain to the other by means of the Fourier Transform.

A variant on the linear adaptive equalizer is the decision feedback equalizer (DFE) [33]. If an equalizer is operating so that the BER is low, then the channel frequency response is known fairly accurately. If so, the transmitted signal can be calculated at the receiver from the received signal and the known frequency response. The echo can then be calculated and the received signal perfectly corrected by subtracting the former from the latter. This method does not add noise as does a linear equalizer. However, to the extent that there are errors in the received signal, this process may increase the error rate. Simple reasoning suggests that there must be a threshold CNR above which the DFE improves the performance and below which it degrades the performance. The crucial situation is at threshold, where the question is whether a DFE extends or diminishes area coverage [40].

No frequency-domain DFE has been reported, but there seems to be no reason why this method could not be used in both systems, if it proved to extend the threshold.

2) Equalization of Dynamic Multipath: Rapidly changing echoes in the presence of a good deal of noise present a serious problem for linear equalizers, since it may not be possible to average over a time long enough to suppress noise in the calculation of equalizer parameters and at the same time follow the dynamic multipath. There seems to be little work reported on this issue. However, a recent paper dealing with MCM indicates that, if the moving echoes are sufficiently random, they may, indeed, be made to add constructively [34]. Presumably, if large fixed echoes could

Table 1 Characteristics of the Three Levels of Performance

Class	Composition		Incremental Rate	Total Rate	Threshold
low-res 384 × 640	MPEG stream audio ancillary data	digital	4 Mb/s	4Mb/s	6 dB CNR
medium-res 576 × 960	enhanced motion vectors selection information additional audio	digital	4 Mb/s	8 Mb/s	17 dB CNR
	selected residual coeffs.	analog	2.5 Ms/s	2.5 Ms/s	
high-res 768 × 1280	enhanced motion vectors selection information additional audio	digital	4 Mb/s	12 Mb/s	29 dB CNR
	selected residual coeffs.	analog	2.5 Ms/s	5 Ms/s	

be made to seem as though they were random, a substantial improvement would result.

D. An Example of a Terrestrial System Having the Desired Properties

We now present the outline of a terrestrial broadcasting system that is "ideal" in the sense that it is intended to meet the requirements previously discussed. It uses some of the techniques that were mentioned earlier and is suitable for use either with a centralized transmitter or in a single-frequency network. The latter gives the highest possible spectrum efficiency; the former gives spectrum efficiency at least as good as the all-digital schemes. It features multiresolution combined source and channel coding. As a result, it supports a good transition scenario and makes possible the manufacture of relatively inexpensive receivers for either configuration of transmitters. Coverage is extended at the lowest performance level and very high resolution is achieved in regions of high signal strength. Interoperability is good, as the signal can easily be decoded at a number of performance levels, the lower levels requiring simpler decoders. Simpler encoders can be used when broadcasting lower-resolution material, such as upconverted NTSC, in which case coverage is further extended. Hybrid analog/digital transmission is used along with a combination of spread spectrum and COFDM for high efficiency and good multipath performance. Digital data is subjected to a powerful forward error-correction process. An all-digital version is available for applications that require it.

The particular system under simulation has a maximum resolution of 768 × 1280 × 60 fps progressively scanned. There are three levels of quality, recoverable at different receiver CNR's, as shown in Table 1. This system is meant to be an example of what can be done with the methods used, and is not a prescription for the best possible scheme for any particular application, although it is thought to be reasonable for use in the US with 6-MHz channels. Fig. 5 shows sample frames at the three levels of resolution. These frames are from a coded sequence with a good deal of motion.

1) Source Coding: A pyramid scheme as in Fig. 1 is used. A high-level block diagram of one level of the coder is shown in Fig. 6. It is clear that the system is closely related to MPEG. The input signal to the coder is the difference between the filtered original and the image as reconstructed by the receiver from the lower levels, if any. A low-pass filter picks out the portion of the difference signal to be coded. The resulting signal is downconverted and the predicted frame at the same level is subtracted. The prediction error is subjected to a wavelet transform (any other transform might be used) and the coefficients to be retained are then adaptively selected. The selected coefficients are transmitted as analog samples and the adaptive selection information is transmitted digitally,[33] using less than one bit/sample.

The predicted frame consists of the previous frame plus a motion-compensated coded version of the predicted change from the last frame to the current frame. Fig. 6 shows the motion estimation being performed by comparing the current frame with the reconstructed previous frame at this coding level. In all likelihood, the final system will calculate the motion vectors directly from the original high-resolution video, using an incremental scheme for the motion information required at each level. Finally, the reconstructed frame is upconverted and subtracted from the input signal to form the input signal for the next level. The decoder at the receiver consists of the elements within the dotted lines.

The lowest level of the pyramid uses MPEG-2 coding and all-digital transmission at a gross data rate of about 10 Mb/s, including audio, forward error correction, and ancillary data. The net corrected video data rate is something less than 4 Mb/s. MPEG coding permits advantage to be taken of available chips. In the simplest receiver, the entire source decoder would consist of a single such chip. The higher levels of the coder generate analog coefficient amplitudes,

[33] This means that the amplitude and identification of the coefficients are not jointly coded, as in MPEG, and that the correlation between these two values is not fully exploited in the compression scheme. Much of this apparent correlation is related to the fact that the selected coefficients are larger and more numerous at lower spatial frequencies and smaller and less numerous at higher spatial frequencies. The sparsity of higher-frequency coefficients is heavily exploited in the vector coder used to transmit the identification of the selected coefficients. The overall efficiency of coding the coefficient information is at least as high as in MPEG

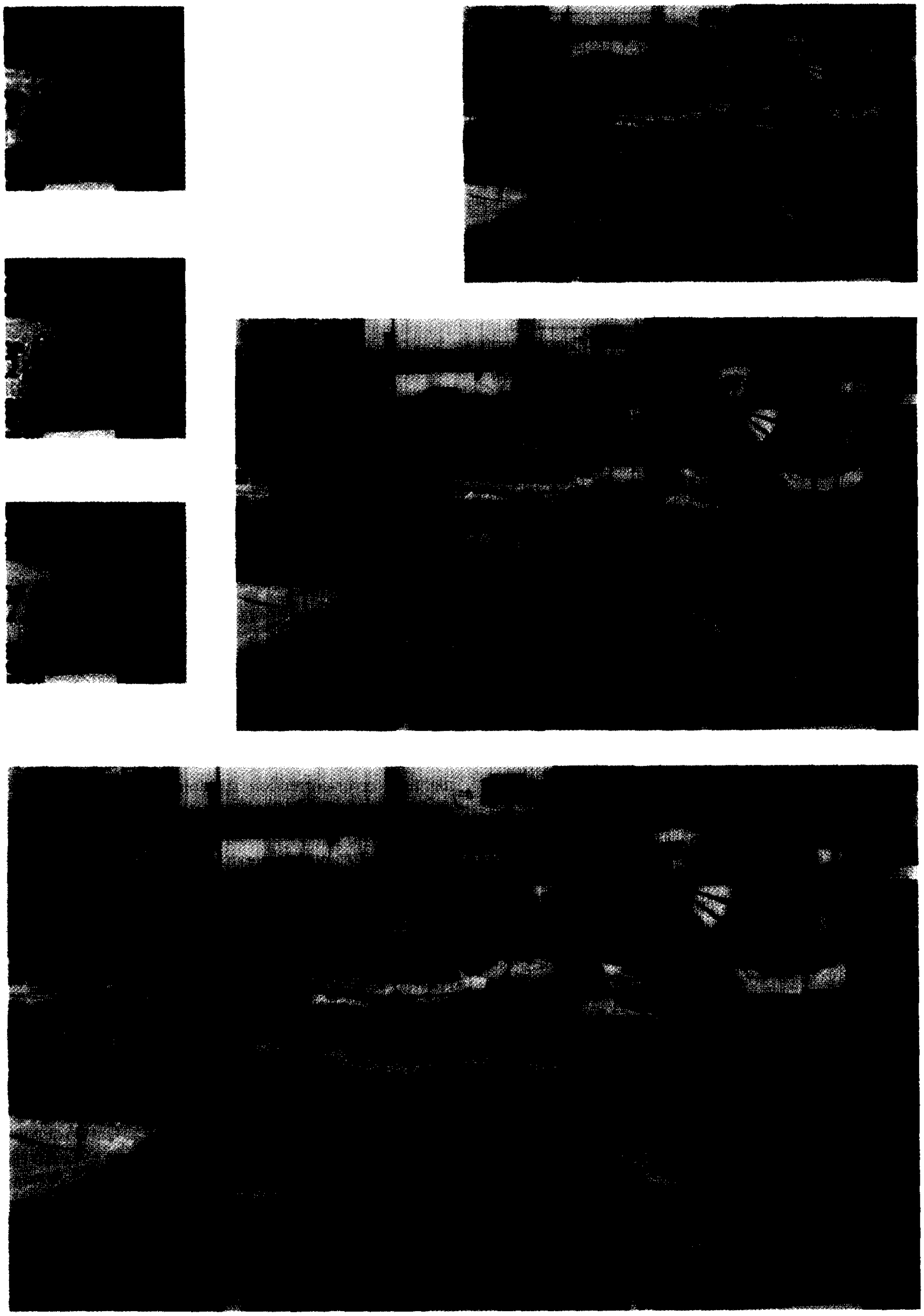


Fig. 5. These three pictures are of a single frame in a sequence with a good deal of motion, produced with the 3-level system described in Section VIII. They are the low-, medium-, and high-resolution versions with the parameters as given in Table 1. All three pictures are somewhat reduced in resolution by the printing process. In order to show the true resolution more accurately, the most difficult portion of each picture has been enlarged to the same size. Defects that may be present in these enlarged pictures are nearly invisible when viewed at normal viewing distance at 60 frames/s.

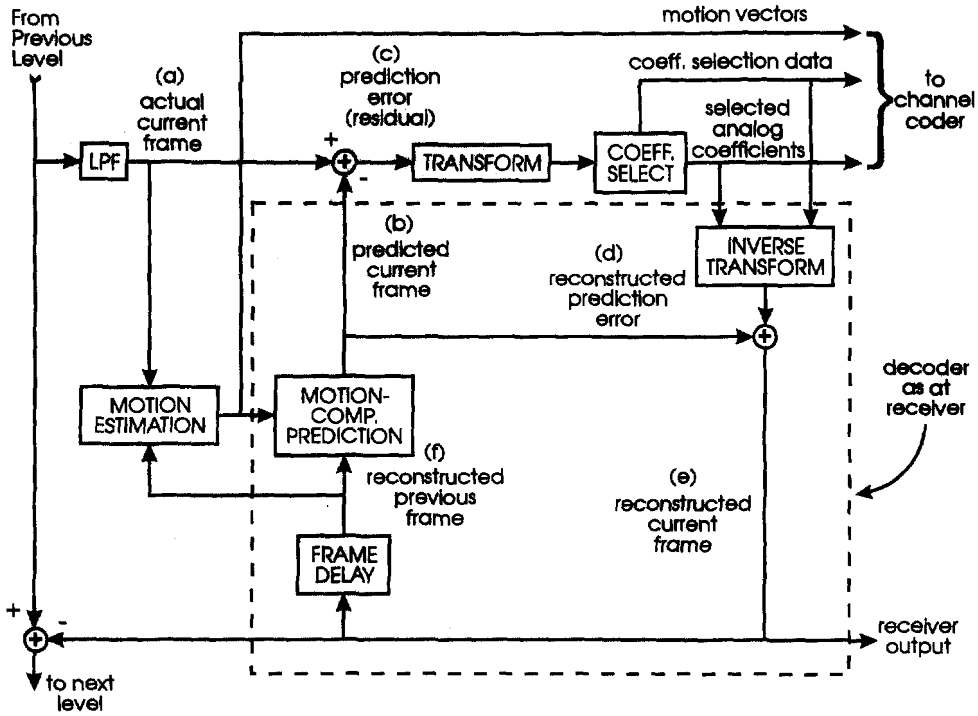


Fig. 6. *One Level of the Pyramid Source Coder/Decoder.* The video data for the current frame that was not coded at lower levels is processed by a low-pass filter, providing the data (a) to be coded at this level. The predicted current frame (b) is then subtracted from the LPF output. The resultant prediction error (c) (the "residual") is subjected to a wavelet transform (any other transform could be used) and the "important" coefficients then selected in a quantity such as not to exceed the allowed transmission capacity, which is 2.5 Msamples/s each for levels 2 and 3. Motion vectors are estimated by comparing the actual current frame with the reconstructed previous frame (f) in the motion estimator. (Motion estimation can be performed in many different ways.) Motion vectors plus coefficient selection data go to the digital input of the channel coder, while selected transform coefficients go to the analog input. Complete decoding for each level, using a method identical to that of the receiver, is required at the encoder in order to produce the reconstructed previous frame and, from it, the predicted current frame, using motion-compensated prediction. The reconstructed current frame (e), which is produced by adding the reconstructed prediction error (d) to the predicted current frame (b), is subtracted from the signal from the previous level to produce the data for the next level, if used. For the purposes of this explanation, it is assumed that there is no delay in any module except the delay module and the motion-compensated predictor. Physical implementation as a pipeline processor requires additional delay modules.

digital coefficient selection data, digital motion vectors, and ancillary data, together with additional audio data, if desired.

Embellishments as used in MPEG and similar systems may, of course, be used here as well. For example, prediction can be bidirectional (at the cost of additional storage) the better to deal with newly revealed areas, a decision between inter- and intraframe coding can be made on a frame-by-frame or block-by-block basis, and the coding can be adapted to the frame rate of the original, as for 24-fps film [35]. On scene changes, the prediction error is naturally much larger than for continuous motion, but the changes can be spread out over several frames to minimize the peak data rate. Of course, scene changes can be flagged.

2) *Channel Coding:* The transmission uses the constellation shown in Fig. 7. It is a nonuniform 64-PSK scheme with 5 Msymbols/s, for a gross data rate of 30 Mb/s[34] and a net error-free data rate of about 12 Mb/s. Digital data sets the angle of the constellation point, and analog data (actually a constant plus the bidirectional coefficient pulse stream at 5 Msamples/s) sets the amplitude. Three digital streams, each of 4 Mb/s, are fed into the three identical error-correction systems, each consisting of an outer (rate .8) Reed–Solomon coder and an inner (rate .5) trellis coder. The output of each of the systems is a four-level (2 bits/sample) stream at 5 Msymbols/s. The three outputs are combined to produce a 64-level signal that determines the angle of the constellation point.

This particular constellation is used because it allows nearly independent decoding of analog and digital data. For the lowest level, with a gross digital data rate of 10 Mb/s and a net error-corrected data rate of only 4 Mb/s, the constellation looks like 4-QAM or 4-PSK, and is very easy to decode. In addition, it is quite robust in the presence of phase noise. It should facilitate the design of less expensive receivers.

The channel coder is shown in Fig. 8. The two streams of analog data from levels two and three of the pyramid coder, each of about 2.5 Ms/s, are weighted, added, and input to the spread-spectrum modulator (SSM). The output of the SSM is an analog data stream at 5 Msamples/s in which each sample is a linear combination of a large number of successive analog coefficients, weighted in such a way that the coefficients of level 2 are recoverable at a lower CNR than those of level 3, and that the relative SNR of the recovered coefficients is optimum according to their spatial frequency. The three streams of digital data are processed

[34] We cannot, of course, expect to receive 30 Mb/s with a usably small error rate except at very high CNR. In order for trellis coding to have a high coding gain at a particular CNR, the equivalent raw error rate of what is sent in the channel must be very high.

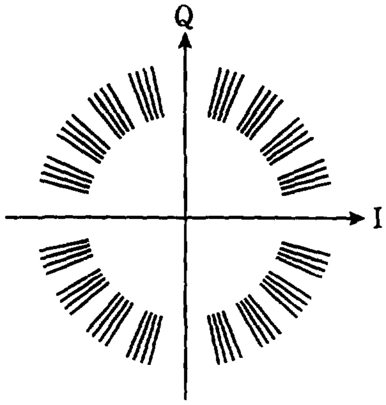


Fig. 7. *The Constellation.* This is the hybrid constellation to be used by the system. Digital data modulates the angle to give nonuniform 64-PSK with about 10–12 dB between levels. The amplitude is a constant plus a function of the analog transform coefficients after spread-spectrum processing. The lengths of the lines are proportional to the rms value of the analog signals.

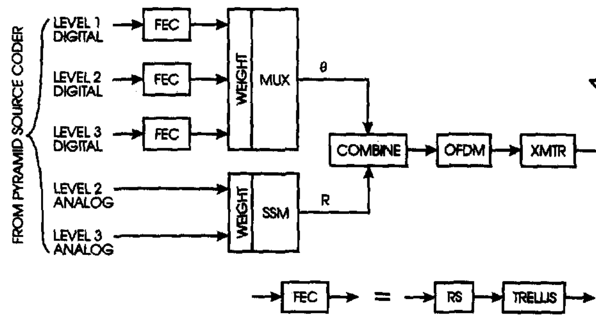


Fig. 8. *Channel Coder.* Digital data from the three levels is processed by three identical forward-error-correction modules, each consisting of a .8 rate Reed–Solomon block coder plus a .5 rate trellis (convolutional) coder. The coded data is weighted and combined in the multiplexer to give the desired angle of the constellation point. Analog data from levels two and three are weighted, subjected to spread-spectrum processing, and added to a constant to produce the desired radial amplitude of the constellation point. The two are combined to produce five complex megasymbols per second. The analog and digital data streams are combined and input to the OFDM processor, whose output goes to the transmitter.

by the FEC as previously described, and then combined with the output of the SSM to form a complex hybrid symbol stream at 5 Msymbols/s. The latter is input to the COFDM processor, which produces a baseband version of the signal for input to the transmitter [36].

The corresponding receiver is shown in Fig. 9. The receiver generates the modulated signal at baseband, corrupted by noise and frequency distortion in the channel. The COFDM demodulator produces a version of the complex hybrid symbol stream, and the properties of the channel (gain, phase, and CNR for each carrier) are estimated on a continuing basis. The amplitude of the demodulated signal is passed to the spread-spectrum demodulator (SSD) along with the channel estimate to produce the coefficients for levels 2 and 3. The phase of the demodulated signal is passed to the demultiplexer, which also makes use of the channel estimates, and is then separated into the three original streams. These are decoded by the error-correction decoders, again using the channel estimates. The recovered analog and digital signals are used in the pyramid decoder to generate the several levels of the video signal.

Two key performance measures for the digital part of the system are shown in Fig. 10. The BER of each of the three data streams, as a function of the CNR in a channel perfect except for noise, is depicted by the solid lines, using the left-hand scale. Notice that the thresholds are separated by 10 to 12 dB. As expected, the performance of each stream is not as good as if that stream had been transmitted by itself, and the performance of all three is limited by the analog data that was added to the digital data. The weighted average SNR of the recovered analog information of the upper two levels is shown in the right-hand scale. Note that the two forms of data are nearly independent, since the phase and amplitude can be decoded separately. The added analog data has some effect on the BER, as does the channel noise.

Note that the thresholds for the three levels of quality are about 6, 17, and 29 dB, when transmitted at full resolution. When transmitting at the lowest resolution only, as for upconverted NTSC, simple 4-PSK is used and the threshold is about 3.2 dB. When transmitting the two lowest levels only, the thresholds are about 5.5 and 15.5 dB.

The dotted lines in Fig. 10 show the performance of both the digital and the analog transmissions in the presence of of echoes. The particular collection of echoes used was one of those used by ATTC in their recent tests of the all-digital systems—the one we judged to be most difficult. Comparison with the solid lines permits an assessment of the degradation of threshold caused by multipath. Note that the quantification of the relative performance of single- and multiple-carrier modulation systems in the presence of multipath is a question that has generated a certain amount of controversy. This measurement is the start of an attempt to answer that question in an empirical manner. The echo results are preliminary.

3) All-Digital Version: For changing this scheme to all-digital, while preserving the maximum similarity so as to enhance interoperability between the digital and hybrid versions, the coefficients need simply to be quantized with an appropriate number of bits/sample and then entropy-coded if desired. Spread spectrum can still be used so as to have two thresholds for the coefficients; the three thresholds for the data that is transmitted digitally in the hybrid version are unchanged. The main effect of using all-digital transmission is that the channel is used less effectively so that somewhat higher CNR is needed in an analog channel for the same picture quality. On the other hand, full digital representation may have some advantages, such as allowing the use of digital VCR's.

IV. Conclusions

We have analyzed the performance factors of an advanced television system for terrestrial broadcasting in the US that are required to maximize its acceptability by the various stakeholders. The latter include regulators,

Table 2 Echo Ensembles Used by ATTC

Ensemble A

Path	Delay	Phase	Atten
1	0.00 μs	288 deg	20 dB
2	1.80 μs	180 deg	0 dB
3	1.95 μs	0 deg	20 dB
4	3.60 μs	72 deg	10 dB
5	7.50 μs	144 deg	14 dB
6	19.80 μs	216 deg	18 dB

Ensemble B

Path	Delay	Phase	Atten
1	0.00 μs	288 deg	20 dB
2	1.75 μs	180 deg	0 dB
3	1.947 μs	0 deg	20 dB
4	3.60 μs	72 deg	10 dB
5	7.50 μs	144 deg	14 dB
6	19.70 μs	216 deg	18 dB

Ensemble C

Path	Delay	Phase	Atten
1	0.00 μs	288 deg	18 dB
2	1.80 μs	180 deg	0 dB
3	1.95 μs	0 deg	20 dB
4	3.60 μs	72 deg	20 dB
5	7.50 μs	144 deg	10 dB
6	19.80 μs	216 deg	14 dB

Ensemble D

Path	Delay	Phase	Attn
1	0.00 μs	288 deg	20 dB
2	1.80 μs	180 deg	0 dB
3	1.95 μs	0 deg	20 dB
4	3.60 μs	72 deg	18 dB
5	7.50 μs	144 deg	14 dB
6	19.80 μs	216 deg	10 dB

Ensemble E

Path	Delay	Phase	Attn
1	0.00 μs	288 deg	20 dB
2	1.80 μs	180 deg	0 dB
3	1.95 μs	0 deg	14 dB
4	3.60 μs	72 deg	10 dB
5	7.50 μs	144 deg	20 dB
6	19.80 μs	216 deg	18 dB

Ensemble F

Path	Delay	Phase	Attn
1	0.00 μs	288 deg	0 dB
2	0.20 μs	180 deg	10 dB
3	1.90 μs	0 deg	14 dB
4	3.90 μs	72 deg	18 dB
5	8.20 μs	144 deg	20 dB
6	15.0 μs	216 deg	20 dB

Ensemble G

Path	Delay	Phase	Attn
1	0.00 μs	180 deg	19 dB
2	0.20 μs	0 deg	0 dB
3	0.28 μs	180 deg	22 dB
4	0.35 μs	180 deg	17 dB
5	0.50 μs	180 deg	22 dB
6	0.80 μs	180 deg	19 dB

These are the seven collections of echoes used in the ATTC tests of the all-digital systems. The Zenith system suffered about a 2.5 dB increase in threshold, averages over all seven collections. The test shown in Fig. 10 used only Collection D, which we judged to be the worst.

broadcasters, equipment manufacturers, program producers, and the viewing public. The factors that emerge as most important are spectrum efficiency, coverage versus quality, cost, interoperability, and the existence of an acceptable transition scenario. As a result of this analysis, we find that existing proposals do not meet all the requirements, and so we have proposed an alternative. The latter makes use of hybrid analog/digital transmission together with joint source and channel coding. It provides several levels of quality according to receiver cost and signal conditions and supports single-frequency operation. A simple receiver can be used for the lowest level of quality, and omnidirectional antennas can be used in most locations.

Appendix

Misconceptions About Digital Broadcasting

Digital processing has many well known advantages over analog processing. For this reason, digital signal processing is already widely used in the TV studio. Digital video tape recorders are now common and, of course, a digital signal representation is needed to utilize these machines. There is also no doubt that digital source coding is superior to analog source coding. For this reason, all the earlier proposed HDTV systems, including MUSE, which uses analog channel coding, use digital source coding. The real issue is whether *all-digital transmission* is required in order to achieve the high compression ratios made possible by digital source coding. The answer is no, as evidenced by the hybrid system described above in Section III-D. Hybrid transmission permits compression comparable to that attainable with digital transmission. At the same time, it permits better utilization of the transmission capacity of the terrestrial broadcasting channel, which, after all, is purely analog. This and other aspects of digital transmission are discussed in the following paragraphs.

A. Utilization of Channel Capacity

This is not an easy subject to address, since there are so many variables and so many differences in the functional characteristics of digital and analog systems. This discussion is, therefore, open to varying interpretations.

An analog HDTV video signal, such as that of the NHK "studio" system, has a bandwidth of about 32 MHz. To fit this within an analog 6-MHz channel requires a bandwidth compression ratio of 5.3. Narrow MUSE attains a ratio of

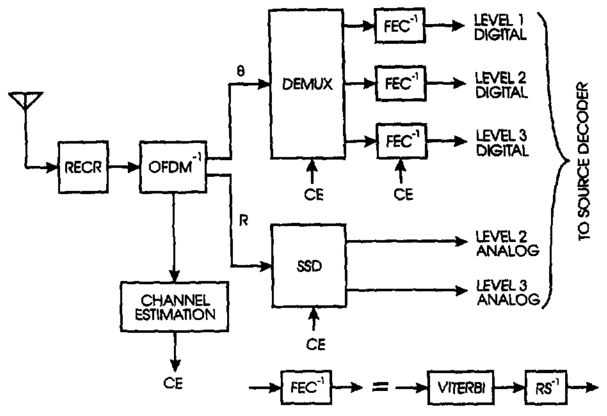


Fig. 9. *Channel Decoder.* The decoder is the inverse of the encoder except for channel equalization. The frequency response of the channel is estimated continuously. The estimate is used in the digital demultiplexer, the error-correction modules, and the spread-spectrum demodulator. The recovered analog and digital data are fed to the source decoder to reconstruct the image.

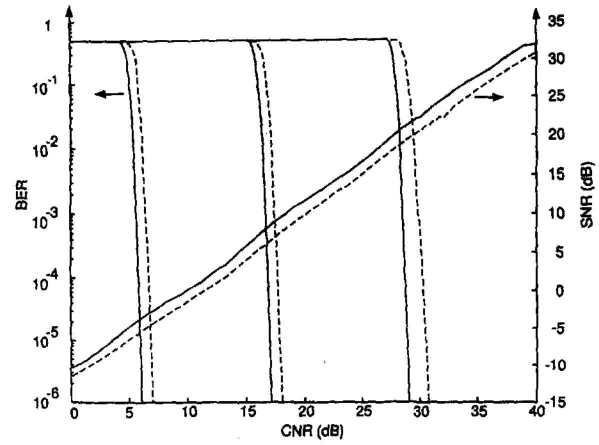


Fig. 10. *Noise Performance, With, and Without Echoes.* Solid lines show the echo-free case while dotted lines show the performance in the presence of echoes from Collection D of Table 2. The BER, after error correction, is shown at the left for each of the three digital data streams, as a function of receiver CNR. The weighted average SNR of the recovered analog signals is shown at the right for the two higher levels. These echo results are preliminary.

4:1 by reduction in diagonal resolution together with a kind of temporal interlace, the latter being made acceptable by motion-adaptive interpolation. The balance of the required compression ratio is achieved by reduction in vertical resolution to 750 lines. Digital systems of comparable picture and sound quality to that of Narrow MUSE, on the other hand, have an uncoded data rate of more than 600 Mb/s, and use about 17 Mb/s for coded video in the channel,[35] for a compression ratio of about 40. Since digital systems are designed to operate with a threshold CNR of about 16 dB, while Narrow MUSE needs about 40 dB, a valid comparison must use a digital channel coder reconfigured to have a threshold of 40 dB. That raises the transmission rate by a factor of 40/16, or 2.5. In that case, the digital source coder would need a compression factor of 16, rather than 40. This can be compared with the value of just 5.3, as needed by an analog system of about the same quality. This comparison between bandwidth compression in an analog system and data compression ratio in a digital system is valid because the noise on the uncompressed analog video has the same effect as channel noise in the kind of coding system used in Narrow MUSE. The ratio 16/5.3 is therefore a measure of the inefficiency of digital transmission in the analog channel. *Thus digital transmission is less, not more, efficient than analog transmission in this case.* Furthermore, at receiving points where the CNR threshold for digital transmission is exceeded, and where the analog system is capable of effective utilization of the additional channel capacity by producing better pictures, the performances of the two kinds of systems diverge even more. Finally, the analog system preserves usable service at CNR's that cause the all-digital schemes to fail entirely.

B. Noise and Interference Rejection

It appears that journalists writing about the "digital revolution" have a vision of distinct ones and zeros (pulses and no pulses) traveling through a channel and being cleaned up by clipping out the noise after reception. Of course, this is not the case in broadcasting. In order to achieve a transmission rate anywhere near the theoretical capacity, large numbers of successive bits must be coded together, complex analog waveforms must be used to represent the blocks of data, and extensive error correction must be used.

Even some of those who do understand the technology persist in making the unqualified statement that digital transmission is more resistant to noise than is analog. This is misleading, since it is only true if the attempted transmission rate is far below the channel capacity. The quantization noise introduced by digital transmission is always larger than the noise that can readily be clipped out. For a valid comparison, the transmission rates of the digital and analog systems must be equal. It has never been proven, and probably is not true, that for a given transmission rate in a channel of given capacity, digital transmission is more resistant to noise than analog.[36]

Noise rejection by clipping[37] is confined to applications in which the transmission rate is well below the channel capacity. In proposed digital cable systems, many programs are to be transmitted on one wire at rates as close to the

[35] For this example, we take a digital system of resolution 720 × 1280 × 60 fps progressively scanned, with the chrominance resolution set at half the luminance resolution in both directions. The compressed data rate is that of the AT&T/Zenith system. The inefficiency comes from many sources, including transmission at less than the Shannon rate, heavy error correction, more audio data, and, perhaps, a less efficient description of the fundamental image information.

[36] When we speak of the "transmission rate" of an analog signal, we must also give an error criterion. A good example would be the case discussed above where a comparison was made between Narrow MUSE and an all-digital system, in which the analog transmission in the noisy channel produced pictures of about the same quality as the digital transmission.

[37] This argument is not confined to simple hard-decision decoders. It applies equally to more sophisticated schemes in which, at the final decision level, a choice is made as to which message was most likely to have been sent, given the received signal and, perhaps, some knowledge of the channel characteristics.

channel capacity as practical and with good error correction. To use repeaters in that case, complete demodulation, decoding to a baseband digital data stream, and recoding would be required at every repeater, a procedure that would be impossibly expensive. In any event, the ability to regenerate digital signals many times in a long series of repeaters with simple reshaping and negligible effect on the BER, which might be applicable to some kinds of long-distance relaying applications, is not relevant to terrestrial broadcasting, where repeaters are not used.

C. Multipath Rejection

One does not see ghosts in digital television pictures, and perhaps this is the reason why some observers have come to believe that digital transmission suppresses ghosts. In fact, the presence of ghosts, even of rather small amplitude, raises the BER to such a degree that digital transmission becomes impossible. *Ghosts must first be removed in order to permit digital transmission at any useful rate.* This is done by some kind of equalization, as discussed in Section III-C. Ironically, should an analog channel be properly equalized, then analog transmission will give greatly improved picture quality. To some extent, this will be done with the "ghost eliminators" that have been developed for NTSC [37].

ACKNOWLEDGMENT

Although this work was supported for the most part by Scitex America, Inc., some MIT funds were also contributed by Prof. Jonathan Allen, Director of the Research Laboratory of Electronics. Graduate students Susie J. Wee and Michael O. Polley did the simulations and also gave material assistance in evolving the the system described in Section III-D and in preparing this paper. Deborah Manning carefully edited the drafts. Valuable comments were made by reviewers, both formal and informal. The errors remaining and the opinions expressed are those of the author alone.

REFERENCES

[1] E. D. Petajan, "Grand alliance television system," this issue.

[2] M. L. Dertouzos *et al., Made in America.* New York: Harper, 1990.

[3] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* Urbana, IL: Univ. Illinois Press, 1963, pp. 67 and 100; T. M. Cover *et al., Elements of Information Theory.* New York: Wiley, 1991.

[4] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE,* vol. PROC–71, pp. 1349–1387, Sept. 1985.

[5] *Digicipher HDTV System Description.* San Diego, CA: General Instrument Corp., 1991.

[6] W. F. Schreiber, "Advanced television systems and their impact on the existing television broadcasting service," submitted to the FCC, MM docket no. 87-286, Nov. 1987, doc. 3, sec. 3.2.

[7] W. R. Neuman, "The mass audience looks at HDTV: An early experiment," in *Nat. Assoc. Broadcast.,* Las Vegas, NV, Apr. 1988.

[8] N. K. Lodge, "The picture quality implications of low bit-rate terrestrial television," in *IEE Colloq. on Develop. in Terrestrial Broadcast. for the UHF Band,* London, Jan. 1992.

[9] O. Bendov, "The effect of channel assignment on transmitter and receiver requirements for equivalent HDTV/NTSC coverage," in *NAB,* Las Vegas, NV, 1994.

[10] Testimony of M. Liebhold of Apple Computer before the House Subcommittee on Tech., Environment, and Aviation, June 1993; J. Sculley, Chairman and CEO, Apple Computer, Nat. Assoc. Broadcast., Las Vegas, Apr. 1993; Computer Syst. Policy Project, letter to J. Quello, May 1993.

[11] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory,* vol. IT-8, p. 2014, 1972.

[12] "Coding of audio, picture, multimedia, and hypermedia information," ISO/IEC cd 13818-2, Dec. 1993, Sec. 7.7–7.9; K. Joseph *et al.,* "MPEG++: A robust compression and transport system for digital HDTV," *Image Commun.,* vol. 4, nos. 4–5, pp. 307–323, Aug. 1992.

[13] W. R. Neuman, private commun.

[14] A. N. Netravali *et al., Digital Pictures.* New York: Plenum, 1988, pp. 497; K. Knowlton, "Progressive transmission of gray-scale and binary pictures by simple, efficient, and lossless encoding schemes," *Proc. IEEE,* vol. PROC–68, pp. 885–896, July 1980.

[15] P. J. Burt *et al.,* "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.,* vol. COM-31, pp. 532–540, 1983; G. Morrison *et al.,* "A spatially layered hierarchical approach to video coding, signal processing," *Image Communication.* New York: Elsevier, 1993, pp. 445–462.

[16] M. Isnardi *et al.,* "A single-channel compatible widescreen EDTV system," 3rd HDTV Colloquium, Canadian Dept. Commun., Ottawa, 1987.

[17] P. G. M. deBot, "Multiresolution transmission over the AWGN channel," Philips Labs., Eindhoven, The Netherlands, June 1992.

[18] "Description of the COFDM System," CCETT, Groupement D'Intéret Economique, Cesson Sevigné, France, 1990; M. Alard and R. Lasalle, "Principles of modulation and channel coding for digital broadcasting to mobile receivers," *EBU Rev.-Tech.,* Aug. 1987, pp. 168–190; P. J. Tourtier *et al.,* "Multicarrier modem for digital HDTV terrestrial broadcasting, signal processing," in *Image Communication.* Amsterdam: Elsevier, vol. 5, 1993, pp. 379–403.

[19] D. Parsons, *The Mobile Radio Propagation Channel.* New York: Wiley, chap. 8.

[20] F. Ferri *et al.,* "Performance analysis of OFDM systems for terrestrial HDTV broadcasting," and M. Kuhn, "The single frequency network of Deutsche Bundepost—Telekom," presented at *Int. Workshop on HDTV,* Torino, Italy, Oct. 1994.

[21] S. A. Lery *et al.,* "Extending HDTV coverage using low power repeaters: A cellular approach," *IEEE Trans. Broadcast.,* vol. 38, pp. 145–150, Sept. 1992.

[22] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.,* vol. COM-19, pp. 628–634, Oct. 1971; B. Hirosaki, "An orthogonally multiplexed QAM system using the discrete Fourier transform," *IEEE Trans. Commun.,* vol. COM-29, pp. 982–989, July 1981.

[23] D. Pommier *et al.,* "A hybrid satellite/terrestrial approach for digital audio broadcast with mobile and portable receivers," in *Proc. NAB Eng. Conf.,* 1990, pp. 304–312.

[24] B. LeFloch, "Orthogonal frequency division multiplex," this issue.

[25] J. N. Ratzel, "The Discrete representation of spatially continuous Images," Sc.D. dissertation, MIT, EECS Dept., 1980; W. F. Schreiber and D. E. Troxel, "Transformation between continuous and discrete representations of images: A perceptual approach," *IEEE Trans. Patt. Anal. and Mach. Intell.,* vol. PAMI-7, pp. 178–186, Feb. 1985.

[26] E. A. Krause, "Motion estimation for frame rate conversion," Ph.D. dissertation, MIT, EECS Dept., 1987; D. M. Martinez, "Model-based motion estimation and its application to restoration and interpolation of motion pictures," Ph.D. dissertation, MIT, EECS Dept., 1986.

[27] G. D. Forney, *Concatenated Codes.* Cambridge, MA: MIT Press, 1966.

[28] A. G. Mason *et al.,* "Digital terrestrial television development in spectre project," IBC, 1992; A. G. Mason *et al.,* "A rugged and flexible digital modulation scheme for terrestrial high definition television," *NAB,* 1992.

[29] M. L. Moher and J. H. Lodge, "A time diversity modulation strategy for the satellite-mobile channel," in *Proc. 13th Biennial Symp. Commun.,* Queen's Univ., Kingston, Canada, June 1986; P. Hoeher, "A statistical discrete-time model for the

WSSUS multipath channel," *IEEE Trans. Vehic. Tech.*, Mar. 1991.

[30] A. J. Viterbi, "Error bounds for convolutional codes and an asymptomatically optimum decoding alogrithm," *IEEE Trans. Inf. Theory,* vol. IT-13, Apr. 1967, pp. 260–269; G. D. Forney, Jr., "The Viterbi algorithm," in *Proc. IEEE,* vol. PROC–61, pp. 268–278, Mar. 1973.

[31] "Test results, grand alliance transmission subsystem," Techn. Subgroup, FCC, ACATS, Jan./Feb. 1994.

[32] T. G. Stockham, Jr., T. M. Cannon, and R. B. Ingebretsen, "Blind deconvolution through digital signal processing," *Proc. IEEE,* vol. PROC–63, pp. 678–692, Apr. 1975.

[33] op. cit, [4].

[34] J. Nicolas *et al.*, "On the performance of multicarrier modulation in a broadcast multipath environment," to appear in *Proc. ICASSP-94.*

[35] W. F. Schreiber *et al.*, "Channel compatible 6-MHz HDTV distribution systems," *SMPTE J.,* pp. 5–13, Jan. 1989.

[36] W. F. Schreiber, "Source and channel coding of advanced TV in the US: Achieving high spectrum efficiency in terrestrial broadcasting," *Münchner Kreis,* Nov. 1993.

[37] "Proposal for ghost cancelling system for NTSC TV," N.A. Philips Labs, May 1991.

[38] H. Hamazumi *et al.*, "Hierarchical TV transmission by spread-spectrum multiplexing," *SMPTE J.,* vol. 103, no. 12, pp. 811–816, Dec. 1994.

[39] W. F. Schreiber, "Spread-spectrum television broadcasting," *SMPTE J.,* vol. 101, no. 8, pp. 538–549, Aug. 1992.

[40] J. J. Nicolas, "Investigation of coding and equalization for the digital HDTV terrestrial broadcast channel," Ph.D. dissertation, EECS Dept., MIT, Cambridge, MA, Sept. 1994. Also Tech. Rep. 585, Res. Lab. of Electron., MIT.

BIBLIOGRAPHY

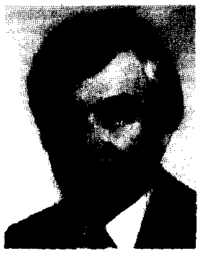
[1] W. F. Schreiber, "Considerations in the design of HDTV systems for terrestrial broadcasting," *SMPTE J.*, vol. 100, no. 9, pp. 668–677, Sept. 1991.

[2] ——, "All-digital HDTV terrestrial broadcasting in the US: Some problems and possible solutions," in *Symp. Int.—Europe-USA,* ENST, Paris, May 1991.

[3] ——, "Advanced television in the United States," Europe. Commission Rep., July 1993.

[4] T. Naveen and J. W. Woods, "Motion compensated multiresolution transmission of high definition video," *IEEE Trans. Circ. and Syst. for Video Tech.,* vol. 4, pp. 29–41, Feb. 1994.

[5] R. Hopkins, "Progress on HDTV broadcasting standards in the United States," in *Signal Processing Image Communication,* vol. 5. New York: Elsevier, 1993, pp. 355–378.



William F. Schreiber (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering from Columbia University. He received the Ph.D. in applied physics from Harvard University in 1953, where he was a Gordon McKay and Charles Coffin fellow.

Dr. Schreiber worked at Sylvania from 1947 to 1949 and at Technicolor Corporation in Hollywood, CA, from 1953 to 1959. From 1959 to 1990, he was a faculty member at MIT, where he is now Professor Emeritus of Electrical Engineering. He was Director of the Advanced Television Research Program from 1983 until his retirement. He was a Visiting Professor of Electrical Engineering at The Indian Institute of Technology, Kanpur, India, during 1964–1966, at INRS-Telecommunications, Montreal, Quebec, during 1981–1982 and 1991–1993, and at the Swiss Federal Institute of Technology, Lausanne, in 1990. He has been interested in image processing since 1948, and has worked in graphic arts, including color processing and laser scanner design, in facsimile, and in television. This work has included theory and extensive practical applications.

Dr. Schreiber is a member of the Technical Association for the Graphic Arts (TAGA), SPIE, and the National Academy of Engineering. He is a fellow of SMPTE. He has received the Honors Award of TAGA, the David Sarnoff Gold Medal from SMPTE, the Gold Medal of the International Society for Optical Engineering (SPIE), and is a four-time recipient of the Journal Award of SMPTE.

A SCALABLE SOURCE CODER FOR A HYBRID HDTV TERRESTRIAL BROADCASTING SYSTEM

Susie J. Wee, Michael O. Polley, and William F. Schreiber

50 Vassar Street, Rm. 36-597
Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139 USA
swee@image.mit.edu

RECEIVED
OCT 1 8 1995
FCC MAIL ROOM

ABSTRACT

A scalable source coder was designed for a hybrid HDTV terrestrial broadcasting system. The hybrid combination of analog and digital methods allows for both high video compression ratios and efficient utilization of the available channel capacity. Efficient channel utilization enables closer receivers to decode high quality video, and allows further receivers to decode natural looking, though lower quality video. The system uses joint source/channel coding to deliver different video components with degrees of integrity reflecting their perceptual importance. Motivation is given for hybrid transmission and analog/digital coding methods. Details of the scalable source coder are provided with a description of its key features, which include pyramid filtering and hybrid video coding. Furthermore, fundamental differences between conventional digital video compression schemes and the proposed scalable hybrid approach are highlighted. Finally, frames of video produced by a computer simulation of the hybrid HDTV system are shown, demonstrating the feasibility of such a system.

1. INTRODUCTION

This paper describes a scalable source coder designed for use in a hybrid[1] (analog/digital) HDTV terrestrial broadcasting system. The system uses joint source/channel coding to transmit different video components with robustness corresponding to their degrees of perceptual importance. The source coder achieves its inherent scalability in video quality, video resolution, and coder/decoder complexity through its two main features: pyramid filtering and hybrid video coding. Pyramid filtering decomposes the video into multiple resolutions of natural looking video, and allows much more coding flexibility than its critically sampled counterparts. Hybrid video coding uses multiresolution motion compensated prediction and adaptive coefficient selection to compress higher resolution enhancement video into hybrid data streams. While many all-digital scalable[2] source coders have been developed in the past [1] [2], to the authors' knowledge this is the first scalable source coder to utilize hybrid methods.

We first describe the concept of hybrid transmission, since it provides motivation for our source coding approach. Next, the significance of joint source/channel coding is discussed. A system description of the scalable source coder is given along with a discussion of various issues concerning scalable hybrid coding. Finally, we show frames of video produced by a computer simulation of the hybrid HDTV terrestrial broadcasting system.

2. HYBRID TRANSMISSION

Conventional HDTV proposals are based on all-digital source coding and transmission; therefore they provide the same video quality level to every receiver in the service area. These schemes are limited in their use of the available channel capacity by the requirements imposed by the receivers operating under the worst channel conditions. Hybrid transmission overcomes this limitation.

Channel Capacity The hybrid HDTV system was designed with the goal of providing the highest possible video quality to each receiver, thereby fully exploiting the channel capacity available to each receiver [3]. In terrestrial broadcasting, the channel capacity available between the central transmitter and each receiver varies, depending on the distance and the channel conditions between the two. Intuitively, a receiver that is close to the transmitter has higher channel capacity than one further away. Therefore, an efficient utilization of the channel should enable closer receivers to reconstruct high quality video, and allow further receivers to reconstruct natural looking, though lower quality video.

Analog/Digital Methods In hybrid transmission, both analog and digital video data are transmitted over the terrestrial broadcast channel. Hybrid transmission allows us to combine the advantages of analog and digital source and channel coding methods. Analog methods are particularly useful in the terrestrial broadcast environment because they allow the quality of the recovered data to vary gracefully as channel conditions vary; this results in efficient channel utilization. Digital methods are particularly useful because reliable reception of digital data allows for the usage of very

This work was supported by Scitex America, Inc.

[1] Throughout this paper, ***hybrid*** refers to the combination of analog and digital data or coding techniques.

[2] multiresolution, hierarchical, layered

powerful video compression techniques. The hybrid combination of analog and digital methods therefore allows for both high video compression ratios and efficient channel utilization.

Multiresolution Coding In the hybrid HDTV system, we allow the quality of the reconstructed video to vary in two ways. First, the quality of the received analog data, and therefore the quality of the reconstructed video, will improve and degrade gracefully with channel conditions. Second, the video may be decoded at several predefined levels of resolution, where the available decodable levels depend on the particular distance and channel conditions between the transmitter and receiver. The hybrid channel coder and scalable source coder described in this paper were designed jointly with the goal of achieving this multiresolution capability.

Channel Coder The hybrid channel coder achieves efficient channel utilization by using multiresolution hybrid modulation. Forward error correction coding, spread-spectrum processing, and orthogonal frequency division multiplexing protect the system against typical channel impairments. A detailed coverage of the hybrid channel coder can be found in [4]. In essence, the hybrid channel coder reliably delivers a digital data stream to all the receivers in the general broadcast area. The general broadcast area is defined as the region in which the received channel signal-to-noise-ratio (CNR) exceeds 6 dB. In addition, a hybrid data stream is decodable where the CNR exceeds 20 dB, and a second hybrid data stream is decodable where the CNR is greater than 32 dB. The signal-to-noise-ratio (SNR) of the received analog portion of the first hybrid data stream is 13 dB where it is first decodable, and increases approximately linearly with CNR; and the analog SNR of the second hybrid data stream is 21 dB where it is first decodable, and also increases approximately linearly with CNR.

3. JOINT SOURCE/CHANNEL CODING

The hybrid HDTV system consists of two parts, the scalable source coder and the hybrid channel coder. The source coder compresses the video signal into multiple analog and digital data streams. The channel coder prepares the data streams for transmission over the broadcast channel. There are many types of video components, each with differing degrees of importance. Therefore, each of the components should be transmitted with a degree of integrity reflecting its importance. For example, motion vectors and low frequency transform coefficients, which have high perceptual importance, should be transmitted with greater robustness than high frequency residual transform coefficients, which have less perceptual importance. Joint source/channel coding is applied to the analog components to distribute the available SNR based on perceptual importance. These tradeoffs are accomplished with a linear transformation, similar to a weighted spread-spectrum operation, which is detailed in [4].

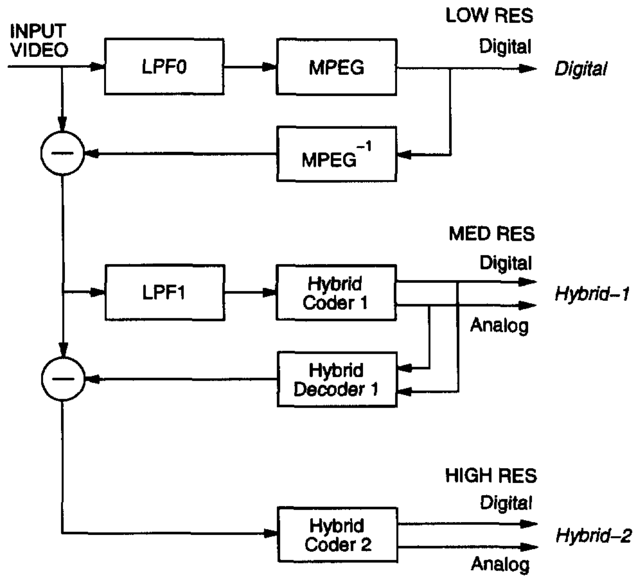


Figure 1: *Scalable Source Coder: The scalable source coder compresses video into three data streams (Digital, Hybrid-1, and Hybrid-2) using pyramid filtering, MPEG video coding, and hybrid video coding. Receivers closest to the transmitter can use all three data streams to decode high resolution video; receivers further from the transmitter can use the Digital and Hybrid-1 data streams to decode medium resolution video; and receivers furthest from the transmitter can use the Digital data stream to reconstruct low resolution video.*

4. SCALABLE SOURCE CODER

The goal of the scalable source coder is to effectively compress video into the three data streams (digital, hybrid-1, and hybrid-2) delivered by the hybrid channel coder. The scalable source coders that have been developed in the past [1] [2] are based on all-digital coding methods; one aspect of the novelty of our approach lies in the use of hybrid coding methods.

As described earlier, the channel coder enables the receivers closest to the transmitter (greater than 32 dB CNR) to decode the one digital and two hybrid data streams; the receivers slightly further from the transmitter (20 to 32 dB CNR) to decode the digital and hybrid-1 data stream; and the receivers furthest from the transmitter (6 to 20 dB CNR) to decode only the digital data stream.

A block diagram of the source coder is shown in Figure 1. Pyramid filtering, conceptually similar to the Laplacian pyramid filtering scheme developed by Burt and Adelson [5], decomposes the video into multiple levels of resolution. An MPEG coder compresses the lowest resolution video into the digital data stream; a hybrid video coder compresses the medium resolution enhancement video into the hybrid-1 data stream; and a second hybrid video coder compresses the high resolution enhancement video into the hybrid-2 data stream. The medium resolution enhancement video is added to the coded low resolution video to form medium resolution video; likewise, the high resolution enhancement video is added to the coded medium resolution video to form high resolution video.

4.1. Pyramid Filtering

The pyramid filtering scheme produces natural looking video at every level of resolution. Its recursive nature provides many advantages and much flexibility over the more commonly used critically sampled schemes. For example, in pyramid filtering, the video coder is given multiple chances to code the low frequency components of the video, which are known to be very important to perceptual video quality. Fewer constraints are placed on the filterbank design, so filters can be chosen for high visual quality. Also, the recursive nature provides flexiblity in the coding that can be performed at each level.

The criteria used in choosing the lowpass filters for each level of the pyramid are visual quality and frequency separation. Sharpened Gaussian filters were chosen for their single overshoot in the spatial domain and their relatively smooth shape and sharp cutoff in the frequency domain. It has been shown that coding with single overshoot filters produces fewer noticeable visual artifacts than multiple overshoot filters [6].

Each successive lowpass filter used in the pyramid is of increasing bandwidth. Therefore, the hybrid coders in the higher pyramid levels encode higher frequency components of the original video. In addition, these coders encode the lower frequency information that was not adequately represented by the lower levels.

Pyramid filtering has been little used because it forms an overcomplete representation of the original image, i.e. the pyramid representation of an image contains more samples than the original image. More popular filtering schemes involve critical sampling, in which the resulting representation contains the same number of samples as the original image. However, we have found that for the purpose of multiresolution coding, the visual quality at each resolution level of the pyramid scheme is superior to that of its critically sampled counterparts.

4.2. Hybrid Video Coding

The hybrid video coder, shown in Figure 2, is used in the higher levels of the pyramid to code the enhancement video into digital and analog data streams. The hybrid video coder is similar to the standard MPEG coder in that motion compensated prediction and a spatial transform are applied. However, rather than using quantization and entropy coding schemes to encode the residual coefficients into a digital data stream, an adaptive coefficient selection technique has been developed to encode the residual coefficients into a hybrid data stream. Also, rather than using the discrete cosine transform, the hybrid video coder uses a wavelet transform that reduces blocking artifacts while maintaining high coding efficiency. The wavelet transform features fine frequency resolution at lower frequencies and fine spatial resolution at higher frequencies; therefore, it is especially well-suited for coding residual video, which is typically highpass and edgy in nature[7].

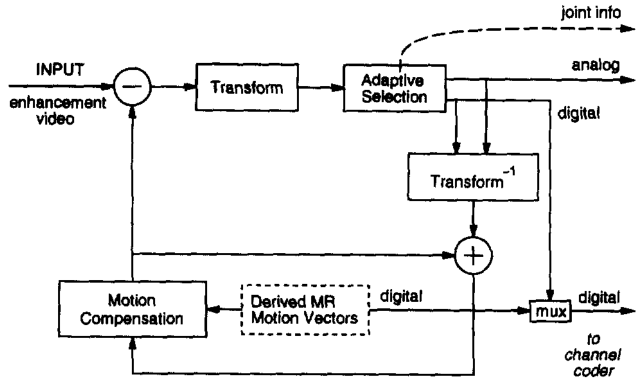


Figure 2: *Hybrid Video Coder: The hybrid video coder compresses enhancement video into a digital and analog data stream, i.e. a hybrid data stream, by using multiresolution motion compensated prediction and adaptive coefficient selection. The motion vectors and the locations of selected coefficients are encoded into the digital data stream; the amplitudes of selected coefficients are communicated in the analog data stream. The received residual coefficients may be noisy, but this noise is preferable to the distortion that results in MPEG-type coders.*

Hybrid Video Coder In the hybrid video coder, a prediction of each video frame is made from previously coded video frames. The prediction is completely specified by motion vectors derived in the encoder and encoded into a digital bit stream. The error in the prediction, called the residual, is transformed into its wavelet representation. A percentage of the residual coefficients is selected for unquantized analog transmission, and the coefficient locations are coded into a digital bit stream. It is assumed that the digital bit stream is received perfectly, and therefore the motion vectors and locations of the residual coefficients are known exactly. The analog signal may suffer from channel noise, and therefore the received residual coefficients may be noisy. However, this noise is preferable to the distortion that results in MPEG-type coders because it is perceptually more similar to the additive noise found in traditional analog video systems.

Adaptive Selection Residual transform coefficients are selected with an adaptive selection process. As previously mentioned, the amplitudes of selected coefficients are communicated by an analog data stream. The locations of selected coefficients are efficiently encoded into a digital bit stream with vector coding methods. The selection is performed in two passes. The coefficients are grouped based on spatial and frequency locations. Initially, groups of coefficients containing a large number of high amplitude coefficients are selected. Final selection is based on local image characteristics. For example, coefficients located at edges of objects are favored over those located in textured areas, such as grass or tree leaves.

Motion Compensated Prediction Motion compensated prediction (MCP) is a highly effective video compression method used in most modern video coding systems. In MCP, each frame of video is predicted from previously coded frames[3]. The prediction is completely specified by a set of motion vectors; accurate knowledge of the motion vectors is essential for decoding video. For this reason, the motion vectors are coded into a digital bit stream. The high coding gain of MCP results because the residual video typically contains much less energy than the original, and is therefore easier to compress. In the few cases where MCP performs badly and the residual is hard to code, for example during scene changes, the original video can be coded instead.

Multiresolution Coding The nature of the pyramid scheme yields redundancy among its different levels. This redundancy is exploited in the multiresolution MCP and adaptive selection methods. Information about the motion vectors at each pyramid level can be derived from those coded in previous levels. Similar relationships can be exploited when coding the locations of coefficients selected at the various levels.

Issues Related to Analog Noise One of the main distinctions between digital and hybrid video coding systems lies in the transmission of the residual coefficients. Digital video coding systems typically transmit quantized residual coefficients, while the hybrid video coder transmits full precision residual coefficients. The digital systems suffer from quantization noise which is known at the encoder, while the hybrid system suffers from channel noise, which is unknown at the encoder. For this reason, it has been speculated that MCP can not be used in a hybrid system. However, test results show that the analog noise performance is good at most analog SNRs in our region of interest. At the very lowest SNRs, the noise does build up after several frames, but leakage or refresh techniques can be used to control this effect. The slight disadvantage in this region is easily outweighed by the graceful performance achieved with hybrid transmission.

5. RESULTS

A very difficult video sequence was processed with a computer simulation of the described hybrid HDTV system. The highly complex sequence was chosen to clearly demonstrate the noise performance of the system. The system performance on a more typical sequence is illustrated in [4].

In the final system, we intend to deliver multiple resolutions of video at 60 frames/sec in a 6 MHz channel with CNR thresholds similar to those quoted earlier. The highest pyramid level will contain 720x1280 resolution video, and each lower level will contain a lower resolution. For practical reasons, our simulation demonstrates a highest resolution of 512x512 pixels, corresponding to a channel bandwidth of 1.7 MHz; a medium resolution of 256x256 pixels; and a lowest resolution of 128x128 pixels. Also, in the final system we intend to provide a more rapid increase in pyramid resolution with thresholds located about 10 dB apart. In addition, a fourth level of resolution, greater than 720x1280 pixels, may be provided.

Figure 3a shows the fifth frame of the original video sequence. Figures 3b-f show the fifth frames of video decoded at various locations in the broadcast area, specifically at 6, 20, 30, 32, and 40 dB CNR. Low resolution video can be decoded in the regions receiving greater than 6 dB CNR. Medium resolution video is first decodable at 20 dB, and video quality improves gracefully with increasing CNR. Improved quality medium resolution video is shown at 30 dB. High resolution video is first decodable at 32 dB, and again, video quality improves gracefully with CNR. Improved quality high resolution video is shown at 40 dB.

[3]This does not imply that successive video frames must be coded consecutively. In fact, if additional coder/decoder complexity and memory requirements are acceptable, then forward, backward, and bidirectional prediction techniques should be employed for improved performance.

6. REFERENCES

[1] K. Uz, M. Vetterli, and D. LeGall, "Interpolative multiresolution coding of advanced television with compatible subchannels," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, March 1991.

[2] T. Naveen and J. Woods, "Motion compensated multiresolution transmission of high definition video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 4, pp. 29–41, February 1994.

[3] W. Schreiber, "Considerations in the design of HDTV systems for terrestrial broadcasting," *SMPTE Journal*, pp. 668–677, September 1991.

[4] M. Polley, S. Wee, and W. Schreiber, "Hybrid channel coding for multiresolution HDTV terrestrial broadcasting," in *ICIP-94*, November 1994.

[5] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Transactions on Communications*, vol. COM-31, pp. 532–540, April 1983.

[6] W. Schreiber and D. Troxel, "Transformation between continuous and discrete representations of images: A perceptual approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, March 1985.

[7] J. Apostolopoulos and J. Lim, "Video compression for digital ATV systems," in *Motion Analysis and Image Sequence Processing* (M. Sezan and R. Lagendijk, eds.), ch. 15, Hingham, MA: Kluwer Academic Publishers, 1993.

Figure 3: a) *Original Video*

Figure 3: b) *Low resolution video: 6 dB CNR*

Figure 3: c) *Medium resolution video: 20 dB CNR*

Figure 3: d) *Medium resolution video: 30 dB CNR*

Figure 3: e) *High resolution video: 32 dB CNR*

Figure 3: f) *High resolution video: 40 dB CNR*

HYBRID CHANNEL CODING FOR MULTIRESOLUTION HDTV TERRESTRIAL BROADCASTING

Michael O. Polley, Susie J. Wee, and William F. Schreiber

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139 USA
email: mpolley@image.mit.edu

ABSTRACT

This paper describes a new HDTV system that applies *joint multiresolution (MR) source and channel coding to* efficiently use the available radio spectrum. Hybrid analog/digital MR channel modulation provides the benefits of digital source coding *and* the more efficient spectrum usage of analog transmission. Nonuniform spacing of the digital signaling levels provides MR delivery of the digital components while spread-spectrum processing permits MR delivery of the analog components. Error correction coding and OFDM channel modulation deliver the MR service in widely varying channel conditions. Simulation results demonstrate performance in various regions of the service area.

1. INTRODUCTION

In this paper we introduce a new HDTV system with particular emphasis on efficient channel coding for terrestrial broadcasting. We approach the terrestrial broadcast problem with the goal of achieving the best possible picture quality at every reception condition [1]. Joint design of hybrid[1] multiresolution (MR) source and channel coders fully utilizes the available channel capacity. The source coder attains its inherent scalability by the use of a pyramid filtering scheme and by the analog transmission of residual coefficients which result from motion compensated prediction applied at each level of the pyramid. A thorough coverage of our MR source coder can be found in [2]. Hybrid MR transmission coupled with spread-spectrum permits reliable delivery of various amounts of coded video information in a scalable fashion, depending on channel conditions. Multicarrier modulation in the form of Orthogonal Frequency Division Multiplexing (OFDM) [3] gives protection against harsh channel conditions and provides an attractive platform to deliver the MR components.

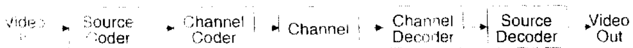


Figure 1: Typical HDTV Terrestrial Broadcast System.

This work was supported by Scitex America, Inc.

[1] Throughout this paper, *hybrid* refers to combined analog and digital data or coding techniques.

As shown in Figure 1, an HDTV broadcasting system contains a source coder to compress high-resolution video signals. The channel coder processes and conditions the coded data for the broadcast channel. The receiver performs the inverse operations to recover the video signal. The conventional approach to designing such a system treats the source coding and channel coding as separate issues. However, this separation in the design is optimal only when the source coder completely eliminates all redundancy in the video signal and the channel coder introduces just enough redundancy to ensure reliable service. Typically these systems use all-digital transmission of a single bit stream with a single threshold of reception, which is not well suited for the terrestrial broadcast environment. Such systems tend to deliver a particular grade of service to all receivers within a certain distance of the transmitter. Receivers within this area get the same picture quality, and those beyond it get nothing at all; only those at threshold utilize the full capacity of the channel.

We consider the design of the source coder and channel coder jointly to achieve better spectrum usage and greater overall coverage. The jointly designed system provides extended coverage at lower quality than single-threshold systems and equal or higher quality in much of the service area. It also features self-optimization at each receiver, depending on signal quality and receiver characteristics.

The following sections give an overview of the system and describe interactions between the source and channel coders. A source coder conducive to hybrid MR transmission is discussed briefly. Then the hybrid MR channel coder and the functionality of each of its components are described. System performance curves describe operation in various broadcast conditions. Finally, a single frame of a video sequence reconstructed at different regions of the service area illustrates system behavior.

2. SYSTEM OVERVIEW

The terrestrial broadcast HDTV system consists of hybrid MR source and channel coders as shown in Figure 2. The channel coder obtains three classes of data from the source coder: low, medium, and high-resolution. The low-resolution data stream contains coded digital video while the medium and high-resolution data streams contain the additional analog and digital data necessary to enhance the video resolu-

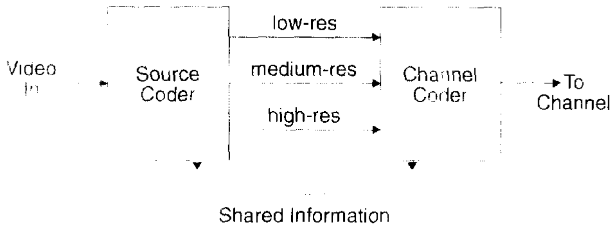


Figure 2: Joint Source and Channel Coder.

tion and quality. The analog enhancement data is composed of selected residual transform coefficients and the digital enhancement data contains location information necessary to use these coefficients. Table 1 lists the composition and net transmission rates for the three data classes.

Class	Composition		Rate
low-res	MPEG stream audio ancillary data	digital	4 Mbps
medium-res	enhanced motion vectors selection information additional audio	digital	4 Mbps
	selected residual coeffs.	analog	2.5 Msps
high-res	enhanced motion vectors selection information additional audio	digital	4 Mbps
	selected residual coeffs.	analog	2.5 Msps

Table 1: Composition and Rates of the MR Data Streams.

The channel coder shares knowledge of the reception thresholds for the three levels of digital data and the overall SNR of the combined analog data streams with the source coder. In return, the source coder shares knowledge of the relative importance of each selected transform coefficient. The source and channel coder jointly determine processing of the analog data so that the most important coefficients are received with highest integrity. As channel conditions improve, the analog coefficients further enhance the reconstructed video.

3 MULTIRESOLUTION SOURCE CODING

The MR source coder features pyramid filtering and hybrid video coding. Pyramid filtering provides scalability by dividing the original video into multiple levels of resolution. Hybrid video coding employs MR motion compensated prediction and adaptive transform coefficient selection techniques to compress video into analog and digital data streams. In the MR source coder shown in Figure 3, low-resolution coded video is formed by low-pass filtering the input video and compressing it to a digital bit stream using a standard MPEG coder. The difference between the original and the MPEG coded low-resolution video, the low-resolution coding error, is processed in the second level with a larger bandwidth low-pass filter and hybrid coder, producing medium-resolution data. The medium-resolution coding error is processed by a second hybrid coder to produce high-resolution data. A three-level source coder is shown here, but any number of levels is possible.

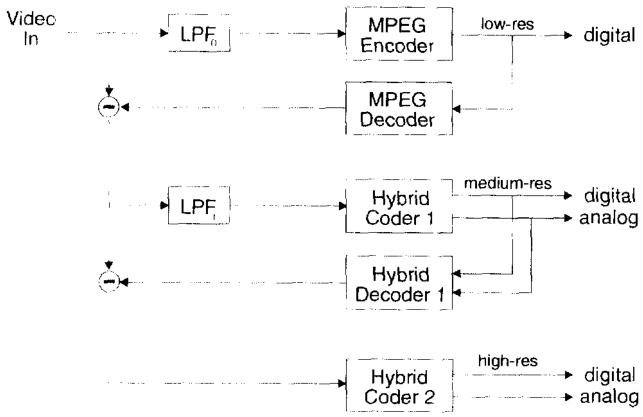


Figure 3: Multiresolution Source Coder.

4. CHANNEL CODING AND TRANSMISSION SYSTEM

Channel coding protects against channel impairments such as random noise, multipath interference (ghosts), and impulse noise in time and frequency. The channel coder achieves robust MR transmission by using a hybrid MR signal constellation, Forward Error Correction (FEC) coding, spread-spectrum processing, and OFDM. The hybrid MR constellation provides a mechanism of combining analog and digital data, and allows retrieval of an increasing number of digital data streams as channel conditions improve. FEC protects the digital bit streams from channel noise and extends the coverage area. Spread-spectrum provides a mechanism for combining analog streams and adjusting the relative quality with which each is recovered. OFDM gives robust transmission in the presence of harsh channel conditions. OFDM coupled with spread-spectrum and FEC techniques make the system resistant to impulse noise in time and frequency. The channel coder processes the different categories of data streams in a manner allowing picture quality to improve as reception conditions improve by using additional MR information to enhance the reconstructed video.

As shown in Figure 4, the channel coder receives multiple data streams from the source coder. Concatenated FEC coding systems, composed of Reed-Solomon (RS) and trellis codes, process the three digital bit streams. The medium and high-resolution analog data streams (transform coefficients) are processed with a spread-spectrum operation using the shared information from the source coder about the relative importance of each datum. The spread-spectrum output and the coded digital data are combined to form a hybrid MR signal that is transmitted using OFDM. At the receiver the inverse coding operations are performed and the various channel impairments are removed.